

Processus de Sauts

Nicolas Broutin

(d'après Philippe Bougerol et Nicolas Fournier)

Master de mathématiques et applications

UE UM4MA236 – Hiver 2026

Version du March 16, 2026

Table des matières

Table des matières	2
1 Marches aléatoires, processus de Bernoulli	5
1.1 Rappels	5
1.2 Marches aléatoires, processus de Bernoulli	6
1.3 Propriété de Markov forte des marches aléatoires	6
1.4 Loi géométrique et processus de Bernoulli	8
1.5 Loi exponentielle	10
2 Processus de Poisson	15
2.1 Processus à accroissements indépendants et stationnaire (P.A.I.S.)	15
2.2 Processus de Poisson	18
2.3 Processus ponctuel de Poisson	21
2.4 La file d'attente $M/G/\infty$	27
3 Processus régénératifs	29
4 Chaînes de Markov	33
4.1 Matrices de transition	33
4.2 Chaîne de Markov	34
4.3 Propriété de Markov	36
4.4 Propriétés de récurrence	38
4.5 Mesures et probabilités invariantes	41
4.6 Réversibilité	47
4.7 Processus de naissance et mort à temps discret	48
4.8 Problèmes d'absorption	50
5 Processus markoviens de sauts	53
5.1 Premières propriétés	53
5.2 Description dynamique	56

5.3	Explosion	59
5.4	Le générateur	61
5.5	Mesures invariantes	64
5.6	Processus de naissance et mort	70
6	Files d'attente et autres applications	73
6.1	Généralités sur les files d'attente	73
6.2	La file $M/M/s$	74
6.3	La file $M/M/1/k$	76
6.4	Interlude sur l'équation de Kolmogorov forward	77
6.5	Un réseau de files d'attentes	77
6.6	Un modèle simple de populations	79
6.7	L'importance des lois exponentielles	81
A	Compléments	83
A.1	π -systèmes et classes monotones	83
A.2	A propos des lois des processus stochastiques	84
A.3	Propriété de Markov pour les trajectoires des chaînes de Markov	86
B	Appendice	89

Marches aléatoires, processus de Bernoulli

Le but de ce chapitre est double. D'une part, faire quelques rappels et fixer les notations. D'autre part, introduire les marches aléatoires et le processus de Bernoulli, caricature à temps discret du très important processus de Poisson.

1.1 Rappels

On se donne un espace de probabilité $(\Omega, \mathcal{F}, \mathbf{P})$.

Soit $(E, \mathcal{E})$ un espace mesurable et X une v.a. à valeurs dans E . La loi de X est la mesure de probabilité μ_X sur $(E, \mathcal{E})$ définie par $\mu_X(A) = \mathbf{P}(X \in A)$, pour tout $A \in \mathcal{E}$.

Pour μ une probabilité sur $(E, \mathcal{E})$, on a $\mu_X = \mu$ si et seulement si $\mathbf{E}[\varphi(X)] = \int_E \varphi d\mu$ pour tout $\varphi \in \mathcal{C}$, où $\mathcal{C}$ est la classe des fonctions mesurables bornées de E dans $\mathbb{R}$. On peut aussi choisir pour $\mathcal{C}$ la classe des fonctions mesurables positives, ou bien encore les fonctions continues bornées.

Définition (Processus stochastique). Soit $T = \mathbb{N}$ ou $\mathbb{R}^+$ (ou éventuellement $\mathbb{Z}$ ou $\mathbb{R}$) et $(E, \mathcal{E})$ un espace mesurable.

- (i) Un processus à valeurs dans E est une famille $(X_t)_{t \in T}$ de v.a. à valeurs dans E .
- (ii) On dit que deux processus $(X_t)_{t \in T}$ et $(Y_t)_{t \in T}$ à valeurs dans E ont la même loi si pour tout $k \geq 1$, pour tout $t_1 < t_2 < \dots < t_k$ (avec $t_i \in T$), les v.a. $(X_{t_1}, \dots, X_{t_k})$ et $(Y_{t_1}, \dots, Y_{t_k})$ (à valeurs dans E^k , qui est muni de la tribu produit $\mathcal{E}^{\otimes k}$) ont la même loi.

L'ensemble T représente généralement le temps, qui peut être discret ou continu.

Exemple 1.1. Attention à ce que l'égalité en loi signifie, notamment lorsque l'intervalle de temps est continu $T = \mathbb{R}_+$: soit U une variable aléatoire uniforme sur $[0, 1]$. On définit $X_t = 0$ pour tout $t \geq 0$ et $\tilde{X}_t = \mathbf{1}_{\{t=U\}}$. Alors pour tout $k \in \mathbb{N}$ et $t_1, t_2, \dots, t_k \in \mathbb{R}_+$, $\mathbf{P}(\exists i \in \{1, 2, \dots, k\} : U = t_i) = 0$, et donc les processus X et $\tilde{X}$ ont même loi! Sans autre condition, le fait que deux processus ont même loi n'implique en aucun cas que leurs trajectoires ont des propriétés similaires.

1.2 Marches aléatoires, processus de Bernoulli

Définition (Marche aléatoire). Soient μ une loi sur $\mathbb{R}^d$ et $X_1, X_2, \dots$ des v.a. i.i.d. de loi μ . Le processus $(S_n)_{n \in \mathbb{N}}$ défini par $S_0 = 0$ et $S_n = X_1 + \dots + X_n$ ($n \geq 1$) est appelé marche aléatoire de “pas” de loi μ .

Exemple 1.2. (i) Si $d = 1$ et si $\mu = (1-p)\delta_0 + p\delta_1$ pour un $p \in [0, 1]$ (i.e. μ est la loi de Bernoulli de paramètre p), alors $(S_n)_{n \in \mathbb{N}}$ est appelé processus de Bernoulli. Bien sûr, pour n fixé, on a $S_n \sim \text{Bin}(n, p)$ et donc $\mathbf{E}[S_n] = np$ et $\mathbf{Var}(S_n) = np(1-p)$. Par exemple, on joue indéfiniment au pile ou face avec une pièce qui tombe sur pile avec probabilité p : S_n représente le nombre de piles obtenus lors des n premiers lancers.

(ii) Si $d = 1$ et si $\mu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, alors le processus $(S_n)_{n \in \mathbb{N}}$ est appelé marche aléatoire simple symétrique.

(iii) Si $d = 2$ et si $\mu = \frac{1}{4}(\delta_{(-1,0)} + \delta_{(0,-1)} + \delta_{(1,0)} + \delta_{(0,1)})$, le processus $(S_n)_{n \in \mathbb{N}}$ est appelé marche aléatoire au plus proche voisin symétrique.

1.3 Propriété de Markov forte des marches aléatoires

La proposition suivante est évidente.

Proposition 1.1 (Propriété de Markov simple pour les M.A.). Soit $(S_n)_{n \in \mathbb{N}}$ une marche aléatoire de pas de loi μ . Pour $n_0 \in \mathbb{N}$ fixé, posons $S_n^{n_0} = S_{n_0+n} - S_{n_0}$. Alors $(S_n^{n_0})_{n \in \mathbb{N}}$ est une marche aléatoire de pas de loi μ et est de plus indépendante de $(S_0, \dots, S_{n_0})$.

Preuve. Il suffit d’observer que

- (i) $S_0^{n_0} = 0$ et $S_{n_0+n} - S_{n_0} = X_{n_0+1} + \dots + X_{n_0+n}$,
- (ii) la suite $(X_{n_0+k})_{k \geq 1}$ est constituée de v.a. i.i.d. de loi commune μ ,
- (ii) la suite $(X_{n_0+k})_{k \geq 1}$ est indépendante du vecteur de variables aléatoires $(X_1, \dots, X_{n_0})$ et donc aussi de $(S_0, S_1, \dots, S_{n_0})$. □

Les définitions qui suivent sont cruciales dans l’étude des processus et seront utilisées tout au long de ce cours.

Définition (Temps d’arrêt). Soit $(S_n)_{n \in \mathbb{N}}$ un processus à temps discret. On introduit la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$ (famille croissante de sous-tribus de $\mathcal{F}$) définie par $\mathcal{F}_n = \sigma(S_0, \dots, S_n)$.

- (i) Une v.a. τ à valeurs dans $\mathbb{N} \cup \{\infty\}$ est un temps d’arrêt (pour le processus $(S_n)_{n \in \mathbb{N}}$) si

$$\forall n \in \mathbb{N}, \quad \{\tau \leq n\} \in \mathcal{F}_n.$$

- (ii) On définit alors la tribu $\mathcal{F}_\tau$ des événements antérieurs à τ par

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq n\} \in \mathcal{F}_n \forall n \in \mathbb{N}\}.$$

- Exercice 1.1.** (i) Une v.a. τ à valeurs dans $\mathbb{N} \cup \{\infty\}$ est un temps d'arrêt ssi $\forall n \in \mathbb{N}, \{\tau = n\} \in \mathcal{F}_n$.
(ii) Montrer que $\mathcal{F}_\tau$ est une tribu
(iii) Montrer que $\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau = n\} \in \mathcal{F}_n \forall n \in \mathbb{N}\}$.

Solution. (i) Supposons que $\{\tau = k\} \in \mathcal{F}_k$ pour tout $k \in \mathbb{N}$. Alors pour $n \in \mathbb{N}$, on a $\{\tau \leq n\} = \bigcup_{0 \leq k \leq n} \{\tau = k\} \in \mathcal{F}_n$. Réciproquement, si $\{\tau \leq k\} \in \mathcal{F}_k$ pour tout $k \in \mathbb{N}$ alors quelque soit $n \in \mathbb{N}$ on a $\{\tau = n\} = \{\tau \leq n\} \cap \{\tau \geq n\} = \{\tau \leq n\} \cap \{\tau \leq n-1\}^c \in \mathcal{F}_n$.

(ii) La seule chose qui n'est pas évidente est la stabilité par passage au complémentaire: pour $A \in \mathcal{F}_\tau$ et $n \in \mathbb{N}$, on écrit $A^c \cap \{\tau \leq n\} = (A \cup \{\tau > n\})^c = ([A \cap \{\tau \leq n\}] \cup \{\tau > n\})^c = [A \cap \{\tau \leq n\}]^c \cap \{\tau \leq n\} \in \mathcal{F}_n$.

(iii) On définit $\mathcal{G}_\tau = \{A \in \mathcal{F} : A \cap \{\tau = n\} \in \mathcal{F}_n \forall n \in \mathbb{N}\}$, de sorte qu'on veut montrer que $\mathcal{G}_\tau = \mathcal{F}_\tau$. Soit $A \in \mathcal{G}_\tau$, alors

$$A \cap \{\tau \leq n\} = \bigcup_{0 \leq k \leq n} \underbrace{A \cap \{\tau = k\}}_{\in \mathcal{F}_k \subset \mathcal{F}_n} \in \mathcal{F}_n$$

et donc $A \in \mathcal{F}_\tau$. De même, pour $A \in \mathcal{F}_\tau$,

$$A \cap \{\tau = n\} = \underbrace{(A \cap \{\tau \leq n\})}_{\in \mathcal{F}_n} \cap \underbrace{(A \cap \{\tau \leq n-1\}^c)}_{\in \mathcal{F}_{n-1} \subset \mathcal{F}_n} \in \mathcal{F}_n,$$

ce qui montre que $A \in \mathcal{G}_\tau$, et donc que $\mathcal{G}_\tau = \mathcal{F}_\tau$. □

Remarque. Le point (i) peut se traduire ainsi : pour tout $n \in \mathbb{N}$, si l'on connaît les valeurs de $S_0, \dots, S_n$, alors on peut dire si $\tau \leq n$.

Pour (iii), un évènement A appartient à $\mathcal{F}_\tau$ si, lorsqu'on connaît les valeurs de $\tau, S_0, \dots, S_\tau$, on peut dire si A est réalisé ou non. Une v.a. Z est $\mathcal{F}_\tau$ -mesurable si, lorsqu'on connaît les valeurs de $\tau, S_0, \dots, S_\tau$, on peut déterminer la valeur de Z .

Exemple 1.3 (Temps d'atteinte). Soit $(S_n)_{n \in \mathbb{N}}$ un processus à valeurs dans $(E, \mathcal{E})$ et $B \in \mathcal{E}$.

(i) Alors $\tau = \inf\{n \geq 0 : S_n \in B\}$, avec la convention $\inf \emptyset = \infty$, est un temps d'arrêt. Si $\tau < \infty$, alors les v.a. $S_\tau, S_{\tau-1}, \sum_{k=0}^{\tau} \varphi(S_k)$ (pour une fonction mesurable $\varphi : E \rightarrow \mathbb{R}$) sont toutes $\mathcal{F}_\tau$ -mesurables. Par contre, $S_{\tau+1}$ n'est en général pas $\mathcal{F}_\tau$ -mesurable.

C'est intuitivement clair : si on fixe $n \in \mathbb{N}$ et qu'on observe $S_0, \dots, S_n$, alors on voit si le processus $(S_k)_{k \in \mathbb{N}}$ est passé dans B avant l'instant n , i.e. on voit si $\tau \leq n$ ou non. Si maintenant on observe $\tau, S_0, S_1, \dots, S_\tau$, i.e. on observe le processus jusqu'à ce qu'il arrive dans B , on peut déterminer les valeurs de $S_\tau, S_{\tau-1}, \sum_{k=0}^{\tau} \varphi(S_k)$, mais pas la valeur de $S_{\tau+1}$.

Formellement, τ est un t.a. car $\forall n \in \mathbb{N}, \{\tau = n\} = \{S_0 \notin B, \dots, S_{n-1} \notin B, S_n \in B\} \in \mathcal{F}_n$. Et par exemple, S_τ est $\mathcal{F}_\tau$ -mesurable car pour tout $\Gamma \in \mathcal{E}, \{S_\tau \in \Gamma\} \in \mathcal{F}_\tau$ car pour tout $n \in \mathbb{N}, \{S_\tau \in \Gamma\} \cap \{\tau = n\} = \{S_0 \notin B, \dots, S_{n-1} \notin B, S_n \in B \cap \Gamma\} \in \mathcal{F}_n$.

(ii) Par contre, $\tau = \sup\{n \geq 0 : S_n \in B\}$ n'est en général pas un temps d'arrêt.

Pour n fixé, en observant $S_0, \dots, S_n$, on ne peut généralement pas dire si $\tau \leq n$, puisqu'on ne sait pas si le processus repassera dans B ou non après n .

Théorème 1.2 (Propriété de Markov forte pour les M.A.). Soit $(S_n)_{n \in \mathbb{N}}$ une marche aléatoire de pas de loi μ . Soit τ un temps d'arrêt (pour $(S_n)_{n \in \mathbb{N}}$ fini p.s. On pose $S_n^\tau = S_{\tau+n} - S_\tau$. Alors $(S_n^\tau)_{n \in \mathbb{N}}$ est une marche aléatoire de pas de loi μ indépendante de $\mathcal{F}_\tau$.

Exemple 1.4. Remarquons qu'il est crucial que τ soit un temps d'arrêt : si par exemple $(S_n)_{n \in \mathbb{N}}$ est un processus de Bernoulli de paramètre $1/2$ et si $\tau = \sup\{n \geq 0 : S_n = 3\}$ (qui n'est pas un temps d'arrêt), alors $S_1^\tau = S_{\tau+1} - S_\tau = 1$ p.s. (sinon $S_{\tau+1} = 3$, ce qui est impossible), donc $(S_n^\tau)_{n \in \mathbb{N}}$ n'est pas un processus de Bernoulli de paramètre $1/2$ (si c'était le cas, on aurait $\mathbf{P}(S_1^\tau = 1) = 1/2$).

Preuve. L'objectif est de montrer que pour tout $A \in \mathcal{F}_\tau$, pour tout $n \geq 0$, pour tous boréliens $B_1, \dots, B_n$ de $\mathbb{R}^d$, on a $I = J$, où

$$I = \mathbf{P}(A, S_1^\tau - S_0^\tau \in B_1, \dots, S_n^\tau - S_{n-1}^\tau \in B_n) \quad \text{et} \quad J = \mathbf{P}(A) \mu(B_1) \dots \mu(B_n).$$

Mais, en se rappelant que $S_n = X_1 + \dots + X_n$ où les X_i sont i.i.d. de loi μ , et en utilisant que τ est p.s. fini,

$$\begin{aligned} I &= \sum_{k \geq 0} \mathbf{P}(A, \tau = k, S_{\tau+1} - S_\tau \in B_1, \dots, S_{\tau+n} - S_{\tau+n-1} \in B_n) \\ &= \sum_{k \geq 0} \mathbf{P}(A, \tau = k, S_{k+1} - S_k \in B_1, \dots, S_{k+n} - S_{k+n-1} \in B_n) \\ &= \sum_{k \geq 0} \mathbf{P}(A, \tau = k, X_{k+1} \in B_1, \dots, X_{k+n} \in B_n). \end{aligned}$$

Comme $A \in \mathcal{F}_\tau$, on a $A \cap \{\tau = k\} \in \mathcal{F}_k$. De plus, $(X_{k+1}, \dots, X_{k+n})$ est indépendant de $\mathcal{F}_k$. Du coup,

$$\begin{aligned} I &= \sum_{k \geq 0} \mathbf{P}(A, \tau = k) \mathbf{P}(X_{k+1} \in B_1, \dots, X_{k+n} \in B_n) \\ &= \sum_{k \geq 0} \mathbf{P}(A, \tau = k) \mu(B_1) \dots \mu(B_n). \end{aligned}$$

Il n'y a plus qu'à remarquer que $\sum_{k \geq 0} \mathbf{P}(A, \tau = k) = \mathbf{P}(A)$ puisque τ est p.s. fini. □

1.4 Loi géométrique et processus de Bernoulli

Considérons un processus de Bernoulli $(S_n)_{n \geq 0}$ de paramètre $p \in]0, 1[$ et

$$T = \inf\{n > 0 : S_n = 1\}.$$

La loi de T s'appelle la loi géométrique (sur $\mathbb{N}^*$) de paramètre p . On a donc

Définition (Loi géométrique). Soit $p \in]0, 1[$. La loi géométrique de paramètre p est la probabilité μ sur $\mathbb{N}^*$ définie par $\mu(\{n\}) = (1-p)^{n-1}p$.

Si X est une v.a. de loi géométrique, on a

$$\mathbf{P}(X > n) = (1-p)^n, \text{ pour tout } n \in \mathbb{N}.$$

Si une expérience aléatoire se répète dans les mêmes conditions, alors le nombre d'essais nécessaires avant le premier succès suit une loi géométrique. Vous jouez au loto: vous attendrez un temps géométrique avant de gagner, etc. Cette loi modélise tous les phénomènes d'attente à temps discret "sans mémoire", au sens où le fait d'attendre beaucoup ne change pas la loi du temps qu'il reste à attendre.

Proposition 1.3 (Absence de mémoire – temps discret). Une v.a. X à valeurs dans $\mathbb{N}^*$ suit une loi géométrique si et seulement si, pour tout $n, m > 0$,

$$\mathbf{P}(X > m + n \mid X > n) = \mathbf{P}(X > m).$$

Preuve. Si X suit une loi géométrique (de paramètre p), alors pour tout $n, m > 0$,

$$\mathbf{P}(X > m + n \mid X > n) = \frac{\mathbf{P}(X > m + n)}{\mathbf{P}(X > n)} = \frac{(1 - p)^{n+m}}{(1 - p)^n} = (1 - p)^m = \mathbf{P}(X > m).$$

Supposons maintenant que pour tout $n, m > 0$, $\mathbf{P}(X > m + n \mid X > n) = \mathbf{P}(X > m)$ et posons $\lambda = \mathbf{P}(X > 1)$. Alors pour tout $n > 0$,

$$\mathbf{P}(X > n + 1) / \mathbf{P}(X > n) = \mathbf{P}(X > n + 1 \mid X > n) = \mathbf{P}(X > 1) = \lambda,$$

d'où

$$\mathbf{P}(X > n + 1) = \lambda \mathbf{P}(X > n).$$

On en déduit que $\mathbf{P}(X > n) = \lambda^n$ pour tout $n > 0$. En particulier, ça implique que $\lambda < 1$ puisque sinon $X = \infty$ p.s. ne serait pas à valeur dans $\mathbb{N}^*$. C'est aussi vrai quand $n = 0$ car $\mathbf{P}(X > 0) = 1$ par hypothèse. Ainsi, pour tout $n > 0$,

$$\mathbf{P}(X = n) = \mathbf{P}(X > n - 1) - \mathbf{P}(X > n) = \lambda^{n-1} - \lambda^n = \lambda^{n-1}(1 - \lambda)$$

et X suit la loi géométrique de paramètre $p = 1 - \lambda$. □

Le résultat suivant est facile à montrer de façon élémentaire. Mais il est plus instructif de le montrer en utilisant la propriété de Markov forte.

Proposition 1.4. Soit $(S_n)_{n \in \mathbb{N}}$ un processus de Bernoulli de paramètre p et $T_k = \inf\{n \geq 0 : S_n = k\}$, pour tout $k \geq 0$ (on a $T_0 = 0$). Alors les variables aléatoires $\{T_{k+1} - T_k, k \geq 0\}$ sont i.i.d. de loi géométrique de paramètre p .

Remarque. On peut écrire $S_n = \sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq n\}}$, car pour tout $\ell \in \mathbb{N}$ fixé,

$$S_n = \ell \iff T_\ell \leq n < T_{\ell+1} \iff \sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq n\}} = \ell.$$

Par la proposition, S_n a donc la même loi que $\sum_{k=1}^{\infty} \mathbf{1}_{\{Y_1 + \dots + Y_k \leq n\}}$, pour une famille $(Y_k)_{k \geq 1}$ de v.a. i.i.d. de loi géométrique de paramètre p .

Exercice 1.2. (i) Si τ est un temps d'arrêt, alors τ est $\mathcal{F}_\tau$ -mesurable (il suffit de montrer que pour tout $k \in \mathbb{N}$, $\{\tau = k\} \in \mathcal{F}_\tau$).

(ii) Si $\sigma \leq \tau$ sont deux temps d'arrêt, alors $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$ (pour $A \in \mathcal{F}_\sigma$ et $n \in \mathbb{N}$, écrire $A \cap \{\tau \leq n\} = (A \cap \{\sigma \leq n\}) \cap \{\tau \leq n\} \in \mathcal{F}_n$).

Preuve de la proposition. C'est un excellent exercice, dont nous ne donnons que les étapes.

- (a) Montrer que T_1 suit la loi géométrique de paramètre p .
- (b) Montrer que pour chaque $k \geq 1$, T_k est un temps d'arrêt (de $(S_n)_{n \in \mathbb{N}}$).
- (c) Montrer que pour chaque $k \geq 1$, $T_1, \dots, T_k$ sont $\mathcal{F}_{T_k}$ -mesurables, et donc que $T_1, T_2 - T_1, \dots, T_k - T_{k-1}$ sont $\mathcal{F}_{T_k}$ -mesurables.
- (d) Fixons $k \geq 1$ et introduisons $\tilde{S}_n = S_n^{T_k} = S_{T_k+n} - S_{T_k}$. Par la propriété de Markov forte, $(\tilde{S}_n)_{n \in \mathbb{N}}$ est un processus de Bernoulli de paramètre p indépendant de $\mathcal{F}_{T_k}$. Donc

$$\tilde{T}_1 = \inf\{n \geq 0 : \tilde{S}_n = 1\}$$

est indépendant de $\mathcal{F}_{T_k}$ et a la même loi que T_1 (la loi géométrique de paramètre p). Mais, comme $S_{T_k} = k$, on a

$$\tilde{T}_1 = \inf\{n \geq 0 : S_{T_k+n} - S_{T_k} = 1\} = \inf\{n \geq 0 : S_{T_k+n} = k+1\} = T_{k+1} - T_k.$$

Ainsi, $T_{k+1} - T_k$ est indépendant de $\mathcal{F}_{T_k}$ et suit la loi géométrique de paramètre p .

- (e) En exploitant les points (c) et (d), on voit que pour tout $k \geq 1$, les v.a. $T_1, T_2 - T_1, \dots, T_{k+1} - T_k$ sont i.i.d. de loi géométrique de paramètre p . C'est ce que nous voulions démontrer. $\square$

1.5 Loi exponentielle

Les lois géométriques sont à valeurs entières. Nous allons considérer maintenant des lois ayant des propriétés analogues, mais à valeurs dans tout $\mathbb{R}^+$: les lois exponentielles. Ce sont, avec les lois de Gauss, les plus importantes du calcul des probabilités à temps continu.

Définition (Loi exponentielle). Soit $\lambda > 0$. On dit qu'une v.a. réelle T suit la loi exponentielle de paramètre λ (en bref, que $T \sim \text{Exp}(\lambda)$) si $\mu_T(dt) = \lambda e^{-\lambda t} \mathbf{1}_{t>0} dt$.

Si $T \sim \text{Exp}(\lambda)$, alors $\mathbf{E}[T] = 1/\lambda$, $\mathbf{Var}(T) = 1/\lambda^2$. Aussi,

$$\mathbf{P}(T \geq t) = \mathbf{P}(T > t) = e^{-\lambda t} \text{ pour tout } t \geq 0.$$

Théorème 1.5 (Absence de mémoire – temps continu). Soit T une v.a. à valeurs dans $\mathbb{R}_*^+$. Les deux énoncés suivants sont équivalents

- (i) il existe $\lambda > 0$ tel que $T \sim \text{Exp}(\lambda)$,
- (ii) pour tout $s, t \geq 0$, $\mathbf{P}(T \geq t+s \mid T \geq t) = \mathbf{P}(T \geq s)$.

C'est très important. (ii) est une propriété qualitative (absence de mémoire). Donc toute v.a. sans mémoire suit nécessairement une loi exponentielle.

Exemple 1.5 (Loi exponentielle et modélisation). (i) Le temps d'attente T , avant la prise d'un premier poisson, d'un pêcheur inexpérimenté est une v.a. sans mémoire (le temps d'attente résiduel ne dépend pas du temps d'attente écoulé). Donc *par nature*, T suit une loi exponentielle. (ii) La durée de vie D d'un objet qui ne s'use pas est une v.a. sans mémoire. Donc *par nature*, D suit une loi exponentielle. Etc.

Preuve. (i) implique (ii), car

$$\mathbf{P}(T \geq t+s \mid T \geq t) = \frac{\mathbf{P}(T \geq t+s)}{\mathbf{P}(T \geq t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = \mathbf{P}(T \geq s).$$

Montrons maintenant que (ii) implique (i). Pour cela, introduisons $G(t) = \mathbf{P}(T \geq t)$. C'est une fonction décroissante sur $\mathbb{R}^+$, on a $G(0) = 1$ et $G(\infty) = 0$ par hypothèse. De plus (ii) donne que $G(t+s) = G(t)G(s)$ pour tous $s, t \geq 0$. On en déduit que pour tous $p, q \in \mathbb{N}$, on a

$$G(p) = G(p/q)^q, \quad G(p) = G(1)^p, \quad \text{d'où} \quad G(p/q) = G(1)^{p/q}.$$

En utilisant que $\mathbb{Q}$ est dense dans $\mathbb{R}^+$, que G est décroissante et que $t \mapsto G(1)^t$ est continue, on en déduit que $G(t) = G(1)^t$ pour tout $t \in \mathbb{R}^+$. De plus,

- $G(1) > 0$, sinon, on aurait $G(t) = 0$ pour tout $t > 0$ et donc $T = 0$ p.s. (or on a supposé T à valeurs dans $\mathbb{R}_*^+$).
- $G(1) < 1$, sinon, on aurait $G(t) = 1$ pour tout $t > 0$, et donc $T = \infty$ p.s. (or on a supposé T à valeurs dans $\mathbb{R}_*^+$).

On pose $\lambda = -\ln G(1) > 0$ et on a $G(t) = e^{-\lambda t}$, i.e. $T \sim \text{Exp}(\lambda)$. □

Lemme 1.6 (des 2 réveils). Soient X_1 et X_2 deux v.a. indépendantes de lois exponentielles de paramètres respectifs λ_1 et λ_2 . Soit $T = \min\{X_1, X_2\}$ et $Z = 1$ si $T = X_1$, $Z = 2$ si $T = X_2$. Alors $T \sim \text{Exp}(\lambda + \mu)$ et est indépendante de Z . De plus $\mathbf{P}(Z = 1) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$ et $\mathbf{P}(Z = 2) = \frac{\lambda_2}{\lambda_1 + \lambda_2}$.

Par exemple, une mare contient des poissons rouges et des poissons gris. Un pêcheur totalement inexpérimenté y pêche. Soit X_1 (resp. X_2) le premier temps de prise d'un poisson rouge (resp. gris). Alors $T = \min\{X_1, X_2\}$ est le temps de prise de son premier poisson, et $Z = 1$ s'il est rouge, 2 s'il est gris. *Par nature* (absence de mémoire), T suit forcément une loi exponentielle. Le lemme des deux réveils, très important, permet de calculer le paramètre de T et la loi de Z .

Prenons l'heure comme unité de temps. Si par exemple $X \sim \text{Exp}(2)$, i.e. on pêche en moyenne 2 poissons rouges par heure (on verra ça plus tard, mais raisonnable car $\mathbf{E}[X_1] = 1/2$), et $X_2 \sim \text{Exp}(3)$, i.e. on pêche en moyenne 3 poissons gris par heure, alors on pêche en moyenne 5 poissons par heure, donc il est raisonnable que $T = \min\{X_1, X_2\} \sim \text{Exp}(5)$, et il est raisonnable de penser que la première prise soit rouge avec probabilité $2/5$.

Preuve. Pour $t > 0$, on a

$$\begin{aligned} \mathbf{P}(T > t, Z = 1) &= \mathbf{P}(X_1 > t, X_1 \leq X_2) \\ &= \int_0^\infty \int_0^\infty \mathbf{1}_{\{x > t, x \leq y\}} \lambda_1 \lambda_2 e^{-\lambda_1 x} e^{-\lambda_2 y} dx dy \\ &= \int_0^\infty \mathbf{1}_{\{x > t\}} \lambda_1 e^{-\lambda_1 x} \left(\int_0^x \lambda_2 e^{-\lambda_2 y} dy \right) dx \\ &= \int_0^\infty \mathbf{1}_{\{x > t\}} \lambda_1 e^{-\lambda_1 x} e^{-\lambda_2 x} dx \\ &= \frac{\lambda_1}{\lambda_1 + \lambda_2} e^{-(\lambda_1 + \lambda_2)t}. \end{aligned}$$

De même,

$$\mathbf{P}(T > t, Z = 2) = \frac{\lambda_2}{\lambda_1 + \lambda_2} e^{-(\lambda_1 + \lambda_2)t}.$$

En sommant, on conclut que $\mathbf{P}(T > t) = e^{-(\lambda_1 + \lambda_2)t}$, donc $T \sim \text{Exp}(\lambda_1 + \lambda_2)$. En prenant $t = 0$, on voit que $\mathbf{P}(Z = 1) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$ et $\mathbf{P}(Z = 2) = \frac{\lambda_2}{\lambda_1 + \lambda_2}$. L'indépendance est aussi démontrée par les formules établies. $\square$

On peut facilement généraliser ce lemme à plus de réveils.

Lemme 1.7 (des réveils). Soit $\{X_k, k = 1, \dots, n\}$ une famille de v.a. indépendantes de lois exponentielles de paramètres $\lambda_1, \dots, \lambda_n$. Soit aussi $\lambda = \sum_{k=1}^n \lambda_k$. Soit

$$T = \min\{X_k : 1 \leq k \leq n\} \quad \text{and} \quad Z = \sum_{k=1}^n k \mathbf{1}_{\{T=X_k\}}.$$

Alors $T \sim \text{Exp}(\lambda)$, T est indépendante de Z , et $\mathbf{P}(Z = k) = \lambda_k / \lambda$ pour tout $k = 1, \dots, n$.

Enfin, rappelons quelques faits sur les lois Gamma. Rappelons d'abord que la fonction Γ d'Euler est définie sur $\mathbb{R}_*^+$ par la formule

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt, \quad x > 0.$$

On vérifie par récurrence, à l'aide d'une intégration par partie, que $\Gamma(n) = (n-1)!$ si $n \in \mathbb{N}^*$.

Définition. On appelle loi Gamma de paramètre (λ, α) où $\lambda, \alpha \geq 0$, la probabilité sur $\mathbb{R}^+$ de densité

$$f_{\lambda, \alpha}(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \mathbf{1}_{\{x > 0\}}.$$

Pour $\alpha = 1$, on retrouve la loi exponentielle.

Proposition 1.8. La somme de deux variables aléatoires indépendantes de lois Gamma de paramètres (λ, α_1) et (λ, α_2) est une variable aléatoire de loi Gamma de paramètre $(\lambda, \alpha_1 + \alpha_2)$.

Preuve. Pour simplifier les notations, on fait la preuve dans le cas où $\lambda = 1$. Pour tout fonction $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ continue bornée, on a

$$\mathbf{E}[\varphi(X + Y)] = \int_0^\infty \int_0^\infty \varphi(x + y) \frac{x^{\alpha_1-1} e^{-x}}{\Gamma(\alpha_1)} \frac{y^{\alpha_2-1} e^{-y}}{\Gamma(\alpha_2)} dx dy.$$

On pose $s = x + y$ and $u = x/(x + y)$ et on effectue le changement de variable $(x, y) \rightarrow (s, u)$ qui est bien un $\mathcal{C}^1$ -difféomorphisme de $]0, \infty[^2$ sur $]0, \infty[\times]0, 1[$ de Jacobien donné par $ds du = s dx dy$ d'où:

$$\begin{aligned} \mathbf{E}[\varphi(X + Y)] &= \int_0^\infty \int_0^1 \varphi(s) s e^{-s} \frac{(su)^{\alpha_1-1} (s(1-u))^{\alpha_2-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} du ds \\ &= \int_0^\infty \varphi(s) \frac{s^{\alpha_1+\alpha_2-1} e^{-s}}{\Gamma(\alpha_1 + \alpha_2)} ds \cdot \int_0^1 u^{\alpha_1-1} (1-u)^{\alpha_2-1} \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} du. \end{aligned}$$

Comme $s^{\alpha_1+\alpha_2-1} e^{-s} / \Gamma(\alpha_1 + \alpha_2) \mathbf{1}_{\{s > 0\}}$ est bien une densité, l'intégrale en u plus haut vaut *forcément* un, ce qui termine la preuve. $\square$

Remarque. Dans la preuve précédente, on a en fait calculé la loi jointe du vecteur $(X + Y, X/(X + Y))$. Si on reprend les calculs avec $E[\varphi(X + Y, X/(X + Y))]$, on voit qu'on a montré que $X/(X + Y)$ était indépendante de $X + Y$ et calculé sa loi: sa densité est la fonction qui apparaît dans l'intégrale en u . Une telle loi est appelée loi Beta de paramètre (α_1, α_2) .

Corollaire 1.9. Si $X_1, X_2, \dots, X_n$ sont des variables aléatoires indépendantes de loi exponentielle de paramètre λ , alors $X_1 + X_2 + \dots + X_n$ a une loi Gamma de paramètre (λ, n) .

Processus de Poisson

2.1 Processus à accroissements indépendants et stationnaire (P.A.I.S.)

Nous voulons généraliser à $\mathbb{R}^+$ la construction de points aléatoires sur $\mathbb{N}$ obtenue par le processus de Bernoulli dans le chapitre précédent. Ceci nous conduit d'abord à la définition suivante, qui généralise au temps continu la notion de marche aléatoire.

Définition (Processus à accroissement indépendants et stationnaires). Un processus $(X_t)_{t \in \mathbb{R}^+}$ à valeurs dans $\mathbb{R}$ est un processus à accroissements indépendants et stationnaires (P.A.I.S.) si, en notant $\mathcal{F}_t = \sigma(X_r, r \leq t)$,

- (i) l'application $t \rightarrow X_t$ est p.s. continue à droite sur $\mathbb{R}^+$;
- (ii) pour tout $s, t \geq 0$, $X_{t+s} - X_s$ est indépendant de $\mathcal{F}_s$;
- (iii) pour tout $s, t \geq 0$, $X_{t+s} - X_s$ a la même loi que $X_t - X_0$;
- (iv) $X_0 = 0$.

Par exemple, le mouvement d'une particule de gaz est souvent modélisée par un P.A.I.S. Plus concrètement, le nombre de poissons N_t pris par un pêcheur inexpérimenté durant $[0, t]$ est un P.A.I.S. (en supposant que la mare contient une infinité de poissons). En effet, les points (i) et (iv) sont satisfaits. De plus, $N_{t+s} - N_t$, qui représente le nombre de poissons pêchés durant $]t, t+s]$ est indépendant de tout ce qui s'est passé jusqu'à l'instant t (point ii) et a bien sûr même loi que le nombre de poissons $N_s = N_s - N_0$ pêchés durant $]0, s]$ (point iii.).

Nous allons généraliser la propriété de Markov forte des marches aléatoires.

Définition (Temps d'arrêt). Soit $(X_t)_{t \in \mathbb{R}_+}$ un processus et, pour $t \geq 0$, $\mathcal{F}_t = \sigma(X_s, s \leq t)$, et $\mathcal{F}_\infty = \sigma(X_s, s \geq 0)$.

- (i) Une v.a. τ à valeurs dans $\mathbb{R}^+ \cup \{\infty\}$ est un temps d'arrêt (pour $(X_t)_{t \in \mathbb{R}_+}$) si

$$\{\tau \leq t\} \in \mathcal{F}_t \quad \text{pour tout } t \geq 0.$$

- (ii) On pose alors $\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty : A \cap \{\tau \leq t\} \in \mathcal{F}_t \forall t \geq 0\}$.

Exercice 2.1. Soit τ et σ deux temps d'arrêt tels que $\tau \leq \sigma$. Montrer que

- (i) $\mathcal{F}_\tau$ est une tribu;
- (ii) $\mathcal{F}_\tau \subseteq \mathcal{F}_\sigma$.

Comme dans le cas discret, un temps aléatoire τ est un temps d'arrêt si pour tout $t \geq 0$, en observant les valeurs de X_s pour $s \in [0, t]$, on est capable de dire si $\tau \leq t$ ou non.

Aussi, pour un événement $A \in \mathcal{F}_\infty$, on a $A \in \mathcal{F}_\tau$ si, lorsqu'on observe τ et X_s pour $s \in [0, \tau]$, on peut dire si A est réalisé ou non. De même, une v.a. Z est $\mathcal{F}_\tau$ -mesurable si l'on peut déterminer sa valeur à partir de celles de τ et des X_s pour $s \in [0, \tau]$.

Lemme 2.1. Tout temps d'arrêt τ est la limite d'une suite décroissante de temps d'arrêt τ_n , où τ_n est à valeurs dans l'ensemble dénombrable $\{k10^{-n}, k \in \mathbb{N}\} \cup \{+\infty\}$.

Preuve. Il suffit de prendre

$$\tau_n = \frac{\lfloor 10^n \tau \rfloor + 1}{10^n}.$$

Alors, τ_n est l'unique élément de $\{10^{-n}k, k \geq 0\}$ tel que $\tau < \tau_n \leq \tau + 10^{-n}$. Par conséquent, comme $10^{-n}\mathbb{N} \subseteq 10^{-(n+1)}\mathbb{N}$, on a $\tau_{n+1} \leq \tau_n$ et τ_n décroît vers τ quand $n \rightarrow \infty$. Pour $n \geq 0$ fixé, montrons que τ_n est un temps d'arrêt. Bien sur, τ_n est $\sigma(\tau)$ -mesurable, et donc $\mathcal{F}_\tau$ -mesurable puisque τ est $\mathcal{F}_\tau$ -mesurable. Donc, comme $\tau_n \geq \tau$, $\{\tau_n \leq t\} = \{\tau_n \leq t\} \cap \{\tau \leq t\}$ est $\mathcal{F}_t$ -mesurable (puisque $\{\tau \leq t\} \in \mathcal{F}_t$). $\square$

Les questions de mesurabilité concernant les temps d'arrêt en temps continu peuvent être délicates. A titre d'exemple montrons le lemme suivant.

Lemme 2.2. Si $(X_t)_{t \in \mathbb{R}^+}$ est continu à droite, pour tout temps d'arrêt τ fini, X_τ est $\mathcal{F}_\tau$ -mesurable.

Preuve. Utilisons le lemme précédent pour écrire $\tau = \lim \tau_n$, où τ_n est une suite décroissante de temps d'arrêt à valeurs dans $\{k10^{-n}, k \in \mathbb{N}\}$. On écrit

$$X_{t \wedge \tau_n} = \sum_{\{k: k10^{-n} \leq t\}} X_{k10^{-n}} \mathbf{1}_{\{\tau_n = k10^{-n}\}} + X_t \mathbf{1}_{\{\tau_n > t\}}.$$

Tous les termes de cette expression sont $\mathcal{F}_t$ -mesurables : pour $s \leq t$, on a $\{\tau_n = s\} \in \mathcal{F}_s \subset \mathcal{F}_t$, X_s est $\mathcal{F}_s$ -mesurable et donc $\mathcal{F}_t$ -mesurable, enfin, $\{\tau_n > t\} = \{\tau_n \leq t\}^c \in \mathcal{F}_t$ puisque τ_n est un temps d'arrêt. Donc $X_{t \wedge \tau_n}$ est $\mathcal{F}_t$ -mesurable.

On en déduit que $X_{\tau \wedge t} = \lim_{n \rightarrow \infty} X_{t \wedge \tau_n}$ est aussi $\mathcal{F}_t$ -mesurable (on utilise ici que X est continu à droite et que la suite τ_n décroît vers τ).

Enfin, pour montrer que X_τ est $\mathcal{F}_\tau$ -mesurable, il faut vérifier que, pour tout borélien B , l'ensemble $\{X_\tau \in B\} \cap \{\tau \leq t\} \in \mathcal{F}_t$. C'est clair car $X_{t \wedge \tau}$ est $\mathcal{F}_t$ -mesurable et

$$\{X_\tau \in B\} \cap \{\tau \leq t\} = \{X_{\tau \wedge t} \in B\} \cap \{\tau \leq t\} \in \mathcal{F}_t. \quad \square$$

Remarque. Comme on vient de le voir, si X est un processus, τ un temps d'arrêt fini, et si $X_{\tau \wedge t}$ est $\mathcal{F}_t$ -mesurable pour tout $t \geq 0$, alors X_τ est aussi $\mathcal{F}_\tau$ -mesurable.

Théorème 2.3 (Propriété de Markov forte des P.A.I.S.). Considérons un P.A.I.S. $(X_t)_{t \in \mathbb{R}^+}$. Pour tout temps d'arrêt τ presque sûrement fini, le processus $(X_{t+\tau} - X_\tau)_{t \in \mathbb{R}^+}$ est de même loi que $(X_t)_{t \in \mathbb{R}^+}$ et est indépendant de $\mathcal{F}_\tau$.

Ce théorème est assez évident si $\tau = t_0$ est déterministe : c'est (en gros) la définition des P.A.I.S.

Preuve. Soit Z une v.a. $\mathcal{F}_\tau$ -mesurable bornée, $\phi_1, \dots, \phi_k$ des fonctions continues bornées sur $\mathbb{R}$ et $0 \leq t_1 < \dots < t_k$. L'objectif est de montrer que $I = J$, où

$$I = \mathbf{E}[Z\phi_1(X_{\tau+t_1} - X_\tau)\phi_2(X_{\tau+t_2} - X_\tau)\cdots\phi_k(X_{\tau+t_k} - X_\tau)]$$

et

$$J = \mathbf{E}[Z]\mathbf{E}[\phi_1(X_{t_1})\phi_2(X_{t_2})\cdots\phi_k(X_{t_k})].$$

Utilisons le lemme 2.1 pour écrire $\tau = \lim \downarrow \tau_n$ où chaque τ_n est à valeurs dans $\{\ell 10^{-n}, \ell \in \mathbb{N}\}$ et introduisons

$$I_n = \mathbf{E}[Z\phi_1(X_{\tau_n+t_1} - X_{\tau_n})\phi_2(X_{\tau_n+t_2} - X_{\tau_n})\cdots\phi_k(X_{\tau_n+t_k} - X_{\tau_n})].$$

Nous allons montrer que $I_n = J$ pour tout n . Cela suffira, puisqu'on a $I = \lim_n I_n$ (car X est continu à droite, car τ_n décroît p.s. vers τ , car les fonctions ϕ_ℓ sont continues, et par convergence dominée, puisque tout est borné). Remarquons que Z est $\mathcal{F}_{\tau_n}$ -mesurable (puisque $\tau \leq \tau_n$) et écrivons

$$\begin{aligned} I_n &= \sum_{\ell \geq 0} \mathbf{E}[Z\mathbf{1}_{\{\tau_n = \ell 10^{-n}\}}\phi_1(X_{\tau_n+t_1} - X_{\tau_n})\cdots\phi_k(X_{\tau_n+t_k} - X_{\tau_n})] \\ &= \sum_{\ell \geq 0} \mathbf{E}[Z\mathbf{1}_{\{\tau_n = \ell 10^{-n}\}}\phi_1(X_{\ell 10^{-n}+t_1} - X_{\ell 10^{-n}})\cdots\phi_k(X_{\ell 10^{-n}+t_k} - X_{\ell 10^{-n}})]. \end{aligned}$$

Pour ℓ fixé, on voit que $Z\mathbf{1}_{\{\tau_n = \ell 10^{-n}\}}$ est $\mathcal{F}_{\ell 10^{-n}}$ -mesurable (car τ_n est un temps d'arrêt et Z est $\mathcal{F}_{\tau_n}$ -mesurable, cf exercice ci-dessous). De plus, par définition des P.A.I.S., le k -uplet

$$(X_{\ell 10^{-n}+t_1} - X_{\ell 10^{-n}}, \dots, X_{\ell 10^{-n}+t_k} - X_{\ell 10^{-n}})$$

est indépendant de $\mathcal{F}_{\ell 10^{-n}}$ et a même loi que $(X_{t_1}, \dots, X_{t_k})$. Ainsi,

$$I_n = \sum_{\ell \geq 0} \mathbf{E}[Z\mathbf{1}_{\{\tau_n = \ell 10^{-n}\}}]\mathbf{E}[\phi_1(X_{t_1})\phi_2(X_{t_2})\cdots\phi_k(X_{t_k})].$$

Cette dernière quantité n'est autre que J , puisque $\sum_{\ell \geq 0} \mathbf{1}_{\{\tau_n = \ell 10^{-n}\}} = 1$. □

Exercice 2.2. Si τ est un t.a. et si Z est $\mathcal{F}_\tau$ -mesurable (à valeurs dans $\mathbb{R}$ par exemple), alors $Z\mathbf{1}_{\{\tau \leq t\}}$ et $Z\mathbf{1}_{\{\tau = t\}}$ sont $\mathcal{F}_t$ -mesurables.

Solution. Pour $Z\mathbf{1}_{\{\tau \leq t\}}$, il s'agit de montrer que pour tout borélien B de $\mathbb{R}$, $\{Z\mathbf{1}_{\{\tau \leq t\}} \in B\} \in \mathcal{F}_t$. Il suffit décrire

$$\{Z\mathbf{1}_{\{\tau \leq t\}} \in B\} = [\{\tau > t\} \cap \{0 \in B\}] \cup [\{\tau \leq t\} \cap \{Z \in B\}] \in \mathcal{F}_t.$$

De manière similaire, on a $\{Z\mathbf{1}_{\{\tau = t\}} \in B\} = (\{0 \in B\} \cap \{\tau \neq t\}) \cup (\{Z \in B\} \cap \{\tau = t\})$. Or $\{\tau \neq t\} = \{\tau < t\} \cup \{\tau \geq t\}^c$ et $\{\tau < t\} = \bigcup_{s < t, s \in \mathbb{Q}} \{\tau \leq s\} \in \mathcal{F}_t$, ce qui montre que $\{0 \in B\} \cap \{\tau \neq t\} \in \mathcal{F}_t$ (puisque $\{0 \in B\}$ est soit $\emptyset$ soit Ω). Finalement,

$$\{Z \in B\} \cap \{\tau = t\} = \underbrace{\{Z \in B\} \cap \{\tau \leq t\}}_{\in \mathcal{F}_t} \cap \underbrace{\{\tau < t\}^c}_{\in \mathcal{F}_t} \in \mathcal{F}_t. \quad \square$$

2.2 Processus de Poisson

Définition (Processus ponctuel). Un processus ponctuel sur $\mathbb{R}^+$ est une suite strictement croissante de v.a. $0 < T_1 < T_2 < \dots$ telle que $T_n \rightarrow \infty$ p.s. quand $n \rightarrow \infty$. Le processus de comptage associé à ce processus ponctuel est le processus $(N_t)_{t \in \mathbb{R}_+}$ défini par

$$N_t = \sum_{k=1}^{+\infty} \mathbf{1}_{\{T_k \leq t\}} = \#\{k \geq 1 : T_k \leq t\}.$$

Remarque. Le processus de comptage $(N_t)_{t \in \mathbb{R}_+}$ de n'importe quel processus ponctuel sur $\mathbb{R}^+$ jouit des propriétés suivantes : $t \rightarrow N_t$ est continu à droite sur $\mathbb{R}^+$, croissant, à valeurs dans $\mathbb{N}$, on a $N_0 = 0$ et $\lim_{t \rightarrow \infty} N_t = \infty$. De plus, pour tout $n \geq 1$, on a

$$T_n \leq t \iff N_t \geq n$$

Enfin, pour tout $n \geq 1$,

$$T_n = \inf\{t \geq 0 : N_t = n\}$$

et T_n est un temps d'arrêt (du processus $(N_t)_{t \in \mathbb{R}_+}$).

Définition (Loi de Poisson). Une v.a. N à valeurs dans $\mathbb{N}$ suit la loi de Poisson de paramètre α , et on note $N \sim \text{Poi}(\alpha)$, si pour tout $n \in \mathbb{N}$,

$$\mathbf{P}(N = n) = e^{-\alpha} \frac{\alpha^n}{n!}.$$

Le résultat suivant est fondamental.

Théorème 2.4. Soit $0 < T_1 < T_2 < \dots$ un processus ponctuel et $(N_t)_{t \geq 0}$ le processus de comptage associé.

(i) Si $(N_t)_{t \geq 0}$ est un P.A.I.S., alors il existe $\lambda > 0$ tel que les v.a. $(T_{n+1} - T_n)_{n \geq 0}$ soient i.i.d. de loi $\text{Exp}(\lambda)$ (avec $T_0 = 0$).

(ii) Soit $\lambda > 0$. Si les v.a. $(T_{n+1} - T_n)_{n \geq 0}$ sont i.i.d. de loi $\mathcal{E}(\lambda)$ (avec $T_0 = 0$), alors $(N_t)_{t \geq 0}$ est un P.A.I.S. De plus, pour tout $0 \leq s < t$, $N_t - N_s \sim \text{Poi}(\lambda(t - s))$.

On dit qu'un tel processus ponctuel $0 \leq T_1 < T_2 < \dots$ (ou, par abus de langage, son processus de comptage $(N_t)_{t \geq 0}$) est un processus de Poisson de paramètre λ .

Par exemple, si N_t est le nombre de poissons pris par un pêcheur inexpérimenté durant $[0, t]$, le processus $(N_t)_{t \geq 0}$ est clairement un P.A.I.S. (on a déjà vu cela). Ainsi, on peut appliquer le théorème.

De même, dans certaines conditions, le nombre de clients N_t arrivant à un péage durant $[0, t]$ est un P.A.I.S. (en gros, les “conditions” reflètent la fluidité du trafic) et on peut appliquer le théorème.

Preuve. Point (i), étape 1 Montrons déjà qu'il existe $\lambda > 0$ tel que $T_1 \sim \mathcal{E}(\lambda)$. Pour cela, il suffit de montrer que pour tout $s, t \geq 0$, $\mathbf{P}(T_1 > t + s) = \mathbf{P}(T_1 > t)\mathbf{P}(T_1 > s)$. Mais

$$\mathbf{P}(T_1 > t + s) = \mathbf{P}(N_{t+s} = 0) = \mathbf{P}(N_{t+s} - N_t = 0, N_t = 0).$$

Comme maintenant $(N_t)_{t \geq 0}$ est un P.A.I.S. par hypothèse, on trouve

$$\mathbf{P}(T_1 > t + s) = \mathbf{P}(N_{t+s} - N_t = 0) \mathbf{P}(N_t = 0) = \mathbf{P}(N_s = 0) \mathbf{P}(N_t = 0),$$

soit encore

$$\mathbf{P}(T_1 > t + s) = \mathbf{P}(T_1 > s) \mathbf{P}(T_1 > t).$$

Point (i), étape 2. Fixons un entier n . Par la propriété de Markov forte du P.A.I.S. $(N_t)_{t \geq 0}$ et comme T_n est un temps d'arrêt, on sait qu'en posant $\tilde{N}_t := N_{T_n+t} - N_{T_n}$, $(\tilde{N}_t)_{t \geq 0}$ a même loi que $(N_t)_{t \geq 0}$ et est indépendant de $\mathcal{F}_{T_n}$. Donc

$$\tilde{T}_1 = \inf\{t \geq 0 : \tilde{N}_t = 1\}$$

a même loi que T_1 (la loi $\mathcal{E}(\lambda)$) et est indépendant de $\mathcal{F}_{T_n}$. Mais, comme $N_{T_n} = n$, on a

$$\tilde{T}_1 = \inf\{t \geq 0 : N_{T_n+t} = n + 1\} = T_{n+1} - T_n.$$

On a montré que pour tout $n \geq 1$, $T_{n+1} - T_n \sim \text{Exp}(\lambda)$ et est indépendant de $\mathcal{F}_{T_n}$. Comme $T_1, \dots, T_n$ sont clairement $\mathcal{F}_{T_n}$ -mesurables, ainsi que $T_1, T_2 - T_1, \dots, T_n - T_{n-1}$, on conclut aisément que les v.a. $(T_{n+1} - T_n)_{n \geq 0}$ sont i.i.d. de loi $\text{Exp}(\lambda)$ (avec $T_0 = 0$).

Point (ii). On va vérifier que pour tout $k \geq 1$, pour tout $0 \leq t_1 < \dots < t_k$, les v.a. $N_{t_i} - N_{t_{i-1}}$ sont indépendantes et de lois $\text{Poi}(\lambda(t_i - t_{i-1}))$. Pour simplifier, on suppose que $k = 3$. On pose $X_1 = T_1$ et, pour $\ell \geq 2$, $X_\ell = T_\ell - T_{\ell-1}$. Par hypothèse, la famille $(X_\ell)_{\ell \geq 1}$ est i.i.d. de loi $\mathcal{E}(\lambda)$. Pour ℓ_1, ℓ_2, ℓ_3 dans $\mathbb{N}$, on a

$$\begin{aligned} I &:= \mathbf{P}(N_{t_1} = \ell_1, N_{t_2} - N_{t_1} = \ell_2, N_{t_3} - N_{t_2} = \ell_3) \\ &= \mathbf{P}\left(T_{\ell_1} \leq t_1 < T_{\ell_1+1}, T_{\ell_1+\ell_2} \leq t_2 < T_{\ell_1+\ell_2+1}, T_{\ell_1+\ell_2+\ell_3} \leq t_3 < T_{\ell_1+\ell_2+\ell_3+1}\right) \\ &= \mathbf{P}\left(X_1 + \dots + X_{\ell_1} \leq t_1 < X_1 + \dots + X_{\ell_1+1}, \right. \\ &\quad \left. X_1 + \dots + X_{\ell_1+\ell_2} \leq t_2 < X_1 + \dots + X_{\ell_1+\ell_2+1}, \right. \\ &\quad \left. X_1 + \dots + X_{\ell_1+\ell_2+\ell_3} \leq t_3 < X_1 + \dots + X_{\ell_1+\ell_2+\ell_3+1}\right) \\ &= \lambda^{\ell_1+\ell_2+\ell_3+1} \int_{\mathbb{R}_+^{\ell_1+\ell_2+\ell_3+1}} e^{-\lambda(x_1+\dots+x_{\ell_1+\ell_2+\ell_3+1})} \\ &\quad \mathbf{1}_{\{x_1+\dots+x_{\ell_1} \leq t_1 < x_1+\dots+x_{\ell_1+1}\}} \mathbf{1}_{\{x_1+\dots+x_{\ell_1+\ell_2} \leq t_2 < x_1+\dots+x_{\ell_1+\ell_2+1}\}} \\ &\quad \mathbf{1}_{\{x_1+\dots+x_{\ell_1+\ell_2+\ell_3} \leq t_3 < x_1+\dots+x_{\ell_1+\ell_2+\ell_3+1}\}} dx_1 \cdots dx_{\ell_1+\ell_2+\ell_3+1}. \end{aligned}$$

On effectue alors le changement de variables triangulaire $s_1 = x_1$, $s_2 = x_1 + x_2$, ..., $s_{\ell_1+\ell_2+\ell_3+1} = x_1 + x_2 + \dots + x_{\ell_1+\ell_2+\ell_3+1}$, dont le jacobien vaut 1, et on trouve

$$\begin{aligned} I &= \lambda^{\ell_1+\ell_2+\ell_3+1} \int_{\mathbb{R}_+^{\ell_1+\ell_2+\ell_3+1}} e^{-\lambda s_{\ell_1+\ell_2+\ell_3+1}} \mathbf{1}_{\{s_1 < s_2 < \dots < s_{\ell_1} \leq t_1\}} \\ &\quad \mathbf{1}_{\{t_1 \leq s_{\ell_1+1} < \dots < s_{\ell_1+\ell_2} \leq t_2\}} \mathbf{1}_{\{t_2 \leq s_{\ell_1+\ell_2+1} < \dots < s_{\ell_1+\ell_2+\ell_3} \leq t_3\}} \\ &\quad \mathbf{1}_{\{s_{\ell_1+\ell_2+\ell_3+1} \geq t_3\}} ds_1 \cdots ds_{\ell_1+\ell_2+\ell_3+1}, \end{aligned}$$

qui se “sépare” bien en $I = \lambda^{\ell_1 + \ell_2 + \ell_3 + 1} \times I_1 \times I_2 \times I_3 \times I_4$, où :

$$\begin{aligned} I_1 &= \int_{(\mathbb{R}^+)^{\ell_1}} \mathbf{1}_{\{s_1 < s_2 < \dots < s_{\ell_1} \leq t_1\}} ds_1 \cdots ds_{\ell_1}, \\ I_2 &= \int_{(\mathbb{R}^+)^{\ell_2}} \mathbf{1}_{\{t_1 \leq s_{\ell_1+1} < \dots < s_{\ell_1+\ell_2} \leq t_2\}} ds_{\ell_1+1} \cdots ds_{\ell_1+\ell_2}, \\ I_3 &= \int_{(\mathbb{R}^+)^{\ell_3}} \mathbf{1}_{\{t_2 \leq s_{\ell_1+\ell_2+1} < \dots < s_{\ell_1+\ell_2+\ell_3} \leq t_3\}} ds_{\ell_1+\ell_2+1} \cdots ds_{\ell_1+\ell_2+\ell_3}, \\ I_4 &= \int_{\mathbb{R}^+} e^{-\lambda s_{\ell_1+\ell_2+\ell_3+1}} \mathbf{1}_{\{s_{\ell_1+\ell_2+\ell_3+1} \geq t_3\}} ds_{\ell_1+\ell_2+\ell_3+1}. \end{aligned}$$

Bien sûr, $I_4 = e^{-\lambda t_3} / \lambda$. On a de plus (exercice), pour tout $k \geq 1$, pour tout $a < b$,

$$J_k(a, b) = \int_{\mathbb{R}^k} \mathbf{1}_{\{a < s_1 < \dots < s_k < b\}} ds_1 \cdots ds_k = \frac{(b-a)^k}{k!}.$$

Comme $I_1 = J_{\ell_1}(0, t_1)$, comme $I_2 = J_{\ell_2}(t_1, t_2)$ et comme $I_3 = J_{\ell_3}(t_2, t_3)$, on trouve

$$I = \lambda^{\ell_1 + \ell_2 + \ell_3} e^{-\lambda t_3} \frac{t_1^{\ell_1}}{\ell_1!} \frac{(t_2 - t_1)^{\ell_2}}{\ell_2!} \frac{(t_3 - t_2)^{\ell_3}}{\ell_3!}.$$

On conclut la preuve en remarquant que

$$I = e^{-\lambda t_1} \frac{(\lambda t_1)^{\ell_1}}{\ell_1!} e^{-\lambda(t_2 - t_1)} \frac{[\lambda(t_2 - t_1)]^{\ell_2}}{\ell_2!} e^{-\lambda(t_3 - t_2)} \frac{[\lambda(t_3 - t_2)]^{\ell_3}}{\ell_3!}. \quad \square$$

Exercice 2.3. Montrer que pour tout entier $k \geq 1$ et tout $a < b$ on a

$$\int_{\mathbb{R}^k} \mathbf{1}_{\{a < x_1 < \dots < x_k < b\}} dx_1, \dots, dx_k = \frac{(b-a)^k}{k!}.$$

Remarque. (i) Soit $(N_t)_{t \geq 0}$ un processus de Poisson de paramètre λ . Alors pour tout $a > 0$, le processus $(N_{t+a} - N_a)_{t \geq 0}$ est un processus de Poisson de paramètre λ (car c’est bien sûr un P.A.I.S. de même loi que $(N_t)_{t \geq 0}$, cela découle directement de la définition des P.A.I.S.). Donc $T_1^a = \inf\{t \geq 0 : N_{t+a} - N_a = 1\}$ suit la loi $\text{Exp}(\lambda)$. Or (exercice) $T_1^a = T_{N_a+1} - a$.

(ii) Paradoxe des autobus. Des autobus arrivent à une station S suivant un processus $(N_t)_{t \geq 0}$ de Poisson de paramètre 4 (l’unité de temps étant l’heure). Il passe donc en moyenne 4 autobus par heure (par exemple, durant la première heure, N_1 autobus sont passés, et comme $N_1 \sim \text{Poi}(4 \times 1)$, $\mathbf{E}[N_1] = 4$). J’arrive à l’instant a . Donc N_a autobus sont déjà passés, le suivant arrivera à l’instant T_{N_a+1} . La durée D de mon attente sera donc égale à $T_{N_a+1} - a$, qui suit une loi $\text{Exp}(4)$, soit en moyenne $\mathbf{E}[D] = 1/4$ d’heure. Comme on arrive “entre” deux autobus et qu’il passe en moyenne un autobus tous les $1/4$ d’heure, on aurait pu penser que la durée moyenne de l’attente serait de $1/8$ heure...

Le dernier résultat concernant les processus de Poisson sur $\mathbb{R}^+$ est tout aussi crucial, c’est une construction alternative à partir de v.a. de loi uniforme.

Proposition 2.5. Soit $T > 0$ et $\lambda > 0$ fixés. Soit $Z \sim \text{Poi}(\lambda T)$ et $(X_i)_{i \geq 1}$ des v.a. i.i.d. de loi uniforme sur $[0, T]$ (indépendantes de Z). Pour $t \in [0, T]$, soit

$$Y_t = \sum_{i=1}^Z \mathbf{1}_{\{X_i \leq t\}}.$$

Alors $(Y_t)_{t \in [0, T]}$ a même loi que $(N_t)_{t \in [0, T]}$, où $(N_t)_{t \geq 0}$ est un processus de Poisson de paramètre λ .

Preuve. Nous allons montrer, et cela suffit, que pour tout $k \geq 1$, pour tout $0 = t_0 < t_1 < \dots < t_{k-1} < t_k = T$, pour tout $n_1 \in \mathbb{N}, \dots, n_k \in \mathbb{N}$, $I = J$, où

$$I = \mathbf{P}(Y_{t_1} = n_1, Y_{t_2} - Y_{t_1} = n_2, \dots, Y_{t_k} - Y_{t_{k-1}} = n_k)$$

et

$$J = \mathbf{P}(N_{t_1} = n_1, N_{t_2} - N_{t_1} = n_2, \dots, N_{t_k} - N_{t_{k-1}} = n_k).$$

On observe que $\{Y_{t_1} = n_1, Y_{t_2} - Y_{t_1} = n_2, \dots, Y_{t_k} - Y_{t_{k-1}} = n_k\}$ ssi $Z = n_1 + \dots + n_k$ et parmi les $n_1 + \dots + n_k$ v.a. X_i , exactement n_1 sont dans $[0, t_1]$, n_2 dans $]t_1, t_2]$, ..., exactement n_k sont dans $]t_{k-1}, t_k]$. Du coup, comme les variables $(X_i)_{i \geq 1}$ sont i.i.d. uniformes sur $[0, T]$, conditionnellement à $Z = n_1 + \dots + n_k$, le vecteur $(N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_k} - N_{t_{k-1}})$ est de loi multinomiale de paramètres $n_1 + n_2 + \dots + n_k$ et $(t_1/T, (t_2 - t_1)/T, \dots, (t_k - t_{k-1})/T)$ et

$$I = e^{-\lambda T} \frac{(\lambda T)^{n_1 + \dots + n_k}}{(n_1 + \dots + n_k)!} \cdot \binom{n_1 + n_2 + \dots + n_k}{n_1, n_2, \dots, n_k} \left(\frac{t_1}{T}\right)^{n_1} \left(\frac{t_2 - t_1}{T}\right)^{n_2} \dots \left(\frac{t_k - t_{k-1}}{T}\right)^{n_k}.$$

En développant tous les coefficients binomiaux, puis en simplifiant, on trouve

$$I = e^{-\lambda T} \lambda^{n_1 + \dots + n_k} \frac{t_1^{n_1}}{n_1!} \frac{(t_2 - t_1)^{n_2}}{n_2!} \dots \frac{(t_k - t_{k-1})^{n_k}}{n_k!}.$$

Enfin, comme $t_1 + (t_2 - t_1) + \dots + (t_k - t_{k-1}) = t_k = T$,

$$I = e^{-\lambda t_1} \frac{(\lambda t_1)^{n_1}}{n_1!} e^{-\lambda(t_2 - t_1)} \frac{[\lambda(t_2 - t_1)]^{n_2}}{n_2!} \times \dots \times e^{-\lambda(t_k - t_{k-1})} \frac{[\lambda(t_k - t_{k-1})]^{n_k}}{n_k!}.$$

C'est bien sûr ce que vaut J , par définition du processus de Poisson. □

2.3 Processus ponctuel de Poisson

Définition (Processus ponctuel de Poisson). Soit $(E, \mathcal{E}, \mu)$ un espace mesuré muni d'une mesure σ -finie μ . Soit $(X_k)_{k \in K}$ un ensemble aléatoire fini ou dénombrable de points de E . Pour $A \in \mathcal{E}$, on pose

$$N_A = \sum_{k \in K} \mathbf{1}_{\{X_k \in A\}}.$$

On dit que $(X_k)_{k \in K}$ est un processus ponctuel de Poisson (PPP) sur E d'intensité μ (et que $(N_A)_{A \in \mathcal{E}}$ est la fonction de comptage associée) si pour tout $n \geq 1$, pour tout $A_1, \dots, A_n \in \mathcal{E}$ deux à deux dis-joints, les v.a. $N_{A_1}, \dots, N_{A_n}$ sont indépendantes et de lois $\text{Poi}(\mu(A_1)), \dots, \text{Poi}(\mu(A_n))$.

Par convention une variable aléatoire de loi de Poisson de paramètre $+\infty$ est une variable aléatoire identiquement infinie.

Remarque. On admettra que la définition ci-dessus caractérise la loi du P.P.P. sur E d'intensité μ . On va montrer l'existence de ce processus quand $\mu(E) < \infty$ dans la Proposition 2.6 et quand $\mu(E) = \infty$ dans la Remarque. Ces énoncés pourront donc être considérés, *a posteriori*, comme des définitions.

Remarquons que N_A n'est autre que le nombre de points (parmi $(X_k)_{k \in K}$) qui appartiennent à A .

Lorsque μ est finie (i.e. $\mu(E) < +\infty$), le nombre total N_E de points du P.P.P. est aléatoire mais fini presque sûrement (puisqu'il suit une loi de Poisson de paramètre $\mu(E)$). Lorsque $\mu(E) = \infty$, N_E est p.s. infini.

Notre premier but est de construire un processus ponctuel de Poisson. Cela prouvera leur existence. Cela permettra aussi de les simuler sur ordinateur. On commence par le cas fini.

Proposition 2.6. Supposons que $\mu(E) < \infty$ et considérons la probabilité $\nu(\cdot) = \mu(\cdot)/\mu(E)$. Donnons-nous des v.a. i.i.d. $U_1, U_2, \dots$ (à valeurs dans E) de loi ν et une v.a. $Z \sim \text{Poi}(\mu(E))$ indépendante de la suite $(U_n)_{n \geq 1}$. Alors

$$\{U_1, U_2, \dots, U_Z\}$$

est un P.P.P. sur E d'intensité μ .

Preuve. Il suffit de montrer que pour $A_1, \dots, A_n$ une partition de E en sous ensembles mesurables, en posant $N_A = \sum_{k=1}^Z \mathbf{1}_{\{U_k \in A\}}$, les v.a. $N_{A_1}, \dots, N_{A_n}$ sont indépendantes de lois de Poisson de paramètres $\mu(A_1), \dots, \mu(A_n)$. Pour $k_1, \dots, k_n \in \mathbb{N}$,

$$\begin{aligned} I &:= \mathbf{P}(N_{A_1} = k_1, N_{A_2} = k_2, \dots, N_{A_n} = k_n) \\ &= \mathbf{P}(Z = k_1 + \dots + k_n, N_{A_1} = k_1, N_{A_2} = k_2, \dots, N_{A_n} = k_n). \end{aligned}$$

Il faut donc que $Z = k_1 + \dots + k_n$ et que parmi les $k_1 + \dots + k_n$ points, k_1 tombent dans A_1 , ..., k_n tombent dans A_n . Comme les variables aléatoires $(U_i)_{i \geq 1}$ sont i.i.d., conditionnellement à $k_1 + \dots + k_n$ le vecteur $(N_{A_1}, \dots, N_{A_n})$ est de loi multinomiale et

$$I = e^{-\mu(E)} \frac{(\mu(E))^{k_1 + \dots + k_n}}{(k_1 + \dots + k_n)!} \cdot \binom{k_1 + k_2 + \dots + k_n}{k_1, k_2, \dots, k_n} (\nu(A_1))^{k_1} (\nu(A_2))^{k_2} \dots (\nu(A_n))^{k_n}.$$

En développant puis en simplifiant les coefficients binômiaux, on trouve

$$I = e^{-\mu(E)} (\mu(E))^{k_1 + \dots + k_n} \frac{(\nu(A_1))^{k_1}}{k_1!} \times \dots \times \frac{(\nu(A_n))^{k_n}}{k_n!}.$$

En se rappelant que $\nu(A) = \mu(A)/\mu(E)$ et que $\mu(E) = \mu(A_1) + \dots + \mu(A_n)$, on conclut que

$$I = e^{-\mu(A_1)} \frac{(\mu(A_1))^{k_1}}{k_1!} \times \dots \times e^{-\mu(A_n)} \frac{(\mu(A_n))^{k_n}}{k_n!},$$

ce qu'on voulait. □

Proposition 2.7 (Principe de superposition). Soit $(\mu_k)_{k \geq 1}$ une famille de mesures σ -finies sur $(E, \mathcal{E})$. Soit $\{(N_A^k)_{A \in \mathcal{E}}, k \in \mathbb{N}\}$, les fonctions de comptage d'une famille de P.P.P. sur E indépendants d'in-

tensités $\{\mu_k, k \in \mathbb{N}\}$. Alors $N_A = \sum_{k=1}^{+\infty} N_A^k$ est la fonction de comptage d'un processus ponctuel de Poisson d'intensité $\mu = \sum_{k \geq 1} \mu_k$.

En termes de processus ponctuels, on en déduit que l'union (dénombrable) de P.P.P. indépendants est un P.P.P. dont l'intensité est la somme des intensités (car sommer les fonctions de comptage revient à faire l'union des ensembles de points).

Remarque. Ceci permet de construire des P.P.P. d'intensité μ σ -finie. En effet, on écrit $\mu = \sum_{k \geq 1} \mu_k$, où chaque mesure μ_k (sur E) est finie, pour chaque k , on construit un P.P.P. en utilisant la proposition 2.6, puis on fait l'union de ces P.P.P. (qu'on a choisis indépendants) en utilisant la propriété de superposition.

Preuve. Il nous faut montrer que si $A_1, \dots, A_n \in \mathcal{E}$ sont deux à deux disjoints, alors $N_{A_1}, \dots, N_{A_n}$ sont indépendants (c'est évident en utilisant l'indépendance des $N_{A_1}^k, \dots, N_{A_n}^k$ pour chaque k et l'indépendance des P.P.P.) et de lois de Poisson de paramètres $\mu(A_1), \dots, \mu(A_n)$: ceci résulte facilement du fait que la somme de v.a. indépendantes de loi de Poisson de paramètres $(\lambda_k)_{k \geq 1}$ suit une loi de Poisson de paramètre $\sum_{k \geq 1} \lambda_k$ (y compris quand $\sum_{k \geq 1} \lambda_k = \infty$, cf exercice). $\square$

Exercice 2.4. Soit X et Y deux variables aléatoires indépendantes telles que $X \sim \text{Poi}(\lambda)$ et $Y \sim \text{Poi}(\mu)$ pour $\lambda, \mu \in \mathbb{R}_+$. Montrer que pour tout $s \geq 0$, $\mathbf{E}[s^X] = e^{\lambda(s-1)}$. En déduire que $X + Y \sim \text{Poi}(\lambda + \mu)$.

Exercice 2.5. Soit $X_1, X_2, \dots$ des v.a. indépendantes de lois de Poisson de paramètres $\lambda_1, \lambda_2, \dots$. On suppose que $\sum_{k \geq 1} \lambda_k = \infty$. Montrer que $S = \sum_{k \geq 1} X_k = \infty$ p.s. On pourra considérer $S_n = \sum_{k=1}^n X_k$, $s_n = \sum_{k=1}^n \lambda_k$, et montrer que $\mathbf{P}(S_n \leq s_n/2) \leq \mathbf{P}(|S_n - s_n| \geq s_n/2) \leq 4\text{Var}(S_n)/s_n^2 = 4/s_n \rightarrow 0$.

Remarque. On note m_+ la mesure de Lebesgue sur $\mathbb{R}^+$. Le processus de Poisson $0 < T_1 < T_2 < \dots$ de paramètre $\lambda > 0$ est un P.P.P. sur $\mathbb{R}^+$ d'intensité λm_+ . Et les notations se transposent ainsi : $N_t = N_{[0,t]}$, $N_t - N_s = N_{]s,t]}$.

Il suffit de comparer la proposition 2.5 et la proposition 2.6.

Un outil souvent utile dans l'étude des processus ponctuels de Poisson est la fonctionnelle de Laplace.

Définition (Fonctionnelle de Laplace). Soit $\{X_n, n \in K\}$ un processus ponctuel sur $(E, \mathcal{E})$. On appelle **fonctionnelle de Laplace** de ce processus l'application $\mathcal{L}$ qui à une fonction mesurable positive $f : E \rightarrow \mathbb{R}^+$ associe

$$\mathcal{L}(f) = \mathbf{E} \left[\exp \left(- \sum_{n \in K} f(X_n) \right) \right].$$

Théorème 2.8. (i) La fonctionnelle de Laplace du processus ponctuel de Poisson sur $(E, \mathcal{E})$ d'intensité μ est donnée par (pour toute $f : E \mapsto \mathbb{R}^+$ mesurable)

$$\mathcal{L}(f) = \exp \left(- \int_E (1 - e^{-f(x)}) \mu(dx) \right).$$

(ii) Réciproquement, si $\{X_n, n \in K\}$ est un processus ponctuel dont la fonctionnelle de Laplace vérifie

$$\mathcal{L}(f) = \exp \left(- \int_E (1 - e^{-f(x)}) \mu(dx) \right)$$

pour toute fonction $f : E \mapsto \mathbb{R}^+$ mesurable prenant un nombre fini de valeurs, alors $\{X_n, n \in K\}$ est un P.P.P. sur $(E, \mathcal{E})$ d'intensité μ .

Preuve. On commence par (i) dans le cas où f ne prend qu'un nombre fini de valeurs (distinctes) $\alpha_1, \dots, \alpha_k$. Si $A_i = \{f = \alpha_i\}$ on a $f = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}$. De plus, les A_i sont deux à deux disjoints par définition. Les v.a. $N_{A_i} = \sum_{n \in K} \mathbf{1}_{\{X_n \in A_i\}}$ sont indépendantes et de loi de Poisson de paramètre $\mu(A_i)$. En se rappelant que si $N \sim \text{Poi}(\lambda)$, alors $\mathbf{E}[\exp(-\alpha N)] = \exp(-\lambda(1 - e^{-\alpha}))$, on trouve donc

$$\begin{aligned} \mathcal{L}(f) &= \mathbf{E} \left[\exp \left(- \sum_{n \in K} \sum_{i=1}^k \alpha_i \mathbf{1}_{\{X_n \in A_i\}} \right) \right] \\ &= \mathbf{E} \left[\exp \left(- \sum_{i=1}^k \sum_{n \in K} \alpha_i \mathbf{1}_{\{X_n \in A_i\}} \right) \right] \\ &= \mathbf{E} \left[\exp \left(- \sum_{i=1}^k \alpha_i N_{A_i} \right) \right] \\ &= \prod_{i=1}^k e^{-\mu(A_i)(1 - e^{-\alpha_i})} \\ &= \exp \left(- \sum_{i=1}^k \mu(A_i)(1 - e^{-\alpha_i}) \right) \\ &= \exp \left(- \int_E (1 - e^{-f(x)}) \mu(dx) \right). \end{aligned}$$

On montre maintenant (ii) : On considère $A_1, \dots, A_k \in \mathcal{E}$ deux à deux disjoints, et on veut montrer que les v.a. $N_{A_1}, \dots, N_{A_k}$ sont indépendantes, de lois de Poisson de paramètres $\mu(A_1), \dots, \mu(A_k)$. Pour cela, il suffit de vérifier que pour tout $\alpha_1, \dots, \alpha_k \geq 0$, on a

$$\mathbf{E} \left[\exp \left(- \sum_{i=1}^k \alpha_i N_{A_i} \right) \right] = \prod_{i=1}^k \exp \left(- \mu(A_i)(1 - e^{-\alpha_i}) \right).$$

C'est précisément ce que dit l'hypothèse appliquée à $f = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}$, qui donne

$$\begin{aligned} \mathbf{E} \left[\exp \left(- \sum_{i=1}^k \alpha_i N_{A_i} \right) \right] &= \mathbf{E} \left[\exp \left(- \sum_{i=1}^k \alpha_i \sum_{n \in K} \mathbf{1}_{\{X_n \in A_i\}} \right) \right] = \mathbf{E} \left[\exp \left(- \sum_{n \in K} f(X_n) \right) \right] \\ &= \mathcal{L}(f) = \exp \left(- \int_E (1 - e^{-f(x)}) \mu(dx) \right) = \prod_{i=1}^k \exp \left(- \mu(A_i)(1 - e^{-\alpha_i}) \right). \end{aligned}$$

On termine maintenant (i) : soit donc f mesurable positive quelconque, et $(f_k)_{k \geq 1}$ une suite de fonctions mesurables positives ne prenant qu'un nombre fini de valeurs croissant vers f (toutes les fonctions

vont de E dans $\mathbb{R}$). Nous savons déjà que pour tout k ,

$$\mathcal{L}(f_k) = \mathbf{E} \left[\exp \left(- \sum_{n \in K} f_k(X_n) \right) \right] = \exp \left(- \int_E (1 - e^{-f_k(x)}) \mu(dx) \right).$$

Il est très facile de conclure en faisant tendre k vers l'infini, par convergences monotone et dominée. $\square$

Le théorème suivant est très important.

Théorème 2.9. Soit $\{X_n, n \in K\}$ un P.P.P. sur E , d'intensité μ et $\{Y_n, n \geq 1\}$ une suite de v.a. i.i.d. à valeurs dans un espace mesurable $(F, \mathcal{F})$, de loi ν , indépendante de $\{X_n, n \geq 1\}$. Alors $\{(X_n, Y_n), n \in K\}$ est un P.P.P. sur $E \times F$ d'intensité $\mu \otimes \nu$.

Le résultat suivant est un corollaire immédiat (choisir $E = \mathbb{R}^+$ et $\mu = \lambda m_+$).

Théorème 2.10 (Processus de Poisson marqué). Soient $0 < T_1 < T_2 < \dots$ les points d'un processus de Poisson sur $\mathbb{R}^+$ de paramètre λ et soit $Y_1, Y_2, \dots$ une suite de variables aléatoires à valeurs dans un ensemble F , i.i.d. de loi ν , et indépendante de $\{T_n, n \geq 1\}$. Alors, $\{(T_n, Y_n), n \geq 1\}$ forme un P.P.P. sur $\mathbb{R}^+ \times F$ d'intensité $\lambda m_+ \otimes \nu$.

Preuve du théorème 2.9. Calculons la fonctionnelle de Laplace du processus $\{(X_n, Y_n), n \in K\}$. Soit $f : E \times F \rightarrow \mathbb{R}$ une fonction mesurable positive. En utilisant l'indépendance des processus $\{X_n, n \in K\}$ et $\{Y_n, n \geq 1\}$, on trouve

$$\begin{aligned} \mathcal{L}(f) &= \mathbf{E} \left[\exp \left(- \sum_{n \in K} f(X_n, Y_n) \right) \right] \\ &= \mathbf{E} \left(\mathbf{E} \left[\exp \left(- \sum_{n \in K} f(X_n, Y_n) \right) \middle| X_k, k \in K \right] \right) \\ &= \mathbf{E} \left(\prod_{n \in K} \mathbf{E} \left[\exp(-f(X_n, Y_n)) \middle| X_k, k \in K \right] \right) \\ &= \mathbf{E} \left(\prod_{n \in K} g(X_n) \right), \end{aligned}$$

où

$$g(x) = \mathbf{E}[e^{-f(x, Y_1)}] = \int_F \exp(-f(x, y)) \nu(dy).$$

Remarquons alors que, puisque $\{X_n, n \in K\}$ est un P.P.P. d'intensité μ ,

$$\mathbf{E} \left(\prod_{n \in K} g(X_n) \right) = \mathbf{E} \left(\exp \left(- \sum_{n \in K} [-\ln g(X_n)] \right) \right) = \exp \left(- \int_E (1 - e^{\ln g(x)}) \mu(dx) \right),$$

soit encore

$$\mathcal{L}(f) = \exp \left(- \int_E (1 - g(x)) \mu(dx) \right).$$

En remplaçant g par sa définition, on obtient donc que

$$\begin{aligned} \mathcal{L}(f) &= \exp \left(- \int_E \left(1 - \int_F e^{-f(x, y)} \nu(dy) \right) \mu(dx) \right) \\ &= \exp \left(- \int_{E \times F} (1 - e^{-f(x, y)}) (\mu \otimes \nu)(dx, dy) \right), \end{aligned}$$

comme désiré. $\square$

Exemple 2.1. Un pêcheur attrape des poissons suivant un processus de Poisson $0 < T_1 < T_2 < \dots$ de paramètre $\lambda > 0$. L'unité de temps est l'heure (il pêche donc en moyenne λ poissons par heure). Notons Y_i la masse du i -ème poisson attrapé. On suppose que les Y_i sont i.i.d. de loi ν (sur $\mathbb{R}^+$) et indépendants du processus de Poisson. On appelle X_t la masse totale de poissons pêchés durant $[0, t]$.

Alors $\{(T_n, Y_n), n \geq 1\}$ est un P.P.P. sur $\mathbb{R}^+ \times \mathbb{R}^+$ d'intensité $\lambda m_+ \otimes \nu$. De plus, $X_t = \sum_{n \geq 1} Y_n \mathbf{1}_{\{T_n \leq t\}} = \sum_{n \geq 1} f(T_n, Y_n)$, où $f(s, y) = y \mathbf{1}_{\{s \leq t\}}$. On peut donc calculer, pour tout $\alpha > 0$,

$$\mathbf{E}[e^{-\alpha X_t}] = \exp\left(-\int_0^\infty \int_0^\infty (1 - e^{-\alpha f(s, y)}) \lambda ds \nu(dy)\right) = \exp\left(-\lambda t \int_0^\infty (1 - e^{-\alpha y}) \nu(dy)\right).$$

Un petit calcul montre enfin que

$$\mathbf{E}[X_t] = -\frac{d}{d\alpha} \mathbf{E}[e^{-\alpha X_t}] \Big|_{\alpha=0} = \lambda t \int_0^\infty y \nu(dy).$$

(On peut aussi le voir directement par linéarité de l'espérance à partir de l'expression de X_t .) C'est logique : durant $[0, t]$, on attrape en moyenne λt poissons, et la masse moyenne des poissons vaut $\int_0^\infty y \nu(dy)$.

Etudions ce qu'on appelle parfois les processus de Poisson effacés.

Corollaire 2.11. Soient $T_1 < T_2 < \dots$, les points d'un processus de Poisson sur $\mathbb{R}^+$ d'intensité λ . Soit $(Y_n)_{n \geq 1}$ une famille i.i.d. de v.a. de Bernoulli de paramètre p (indépendante du processus de Poisson). Alors

$$N_t^0 = \sum_{n \geq 1} (1 - Y_n) \mathbf{1}_{\{T_n \leq t\}} \quad \text{et} \quad N_t^1 = \sum_{n \geq 1} Y_n \mathbf{1}_{\{T_n \leq t\}}$$

sont les processus de comptage de deux processus de Poisson indépendants de paramètres respectifs $\lambda(1-p)$ et λp .

Preuve. On sait que $\{(T_n, Y_n), n \geq 1\}$ est un P.P.P. sur $\mathbb{R}^+ \times \{0, 1\}$ d'intensité $\xi = \lambda m_+ \otimes ((1-p)\delta_0 + p\delta_1)$. Soit N sa fonction de comptage. On a alors, pour tout $0 \leq s < t$ (noter que $1 - Y_n = \mathbf{1}_{\{Y_n=0\}}$ et $Y_n = \mathbf{1}_{\{Y_n=1\}}$)

$$N_t^0 - N_s^0 = N([s, t] \times \{0\}) \quad \text{et} \quad N_t^1 - N_s^1 = N([s, t] \times \{1\}).$$

L'indépendance de N^0 et N^1 s'ensuit facilement : pour tout $0 < s < t$ et $0 < u < v$, les ensembles $]s, t] \times \{0\}$ et $]u, v] \times \{1\}$ sont disjoints. Pour montrer que N^1 (par exemple) est un processus de Poisson de paramètre λp , il suffit de montrer que pour tout $0 = t_0 < t_1 < \dots < t_k$, les v.a. $N_{t_i}^1 - N_{t_{i-1}}^1, i = 1, \dots, k$ sont indépendantes de lois de Poisson de paramètre $\lambda p(t_i - t_{i-1})$. L'indépendance découle du fait que les ensembles $]t_{i-1}, t_i] \times \{1\}$ sont deux à deux disjoints. De plus, on sait que $N_{t_i}^1 - N_{t_{i-1}}^1 = N([t_{i-1}, t_i] \times \{1\})$ suit une loi de Poisson de paramètre $\xi([t_{i-1}, t_i] \times \{1\})$, qui vaut bien $\lambda p(t_i - t_{i-1})$. $\square$

Exercice 2.6. Un poisson pond N oeufs, où N suit une loi de Poisson de paramètre λ . Chaque oeuf éclot avec probabilité $p \in]0, 1[$, indépendamment. Soit X le nombre de bébés poissons. Montrer que X suit la loi de Poisson de paramètre λp . C'est bien sûr très lié au corollaire précédent.

Exemple 2.2. Un pêcheur attrape des poissons suivant un processus de Poisson $0 < T_1 < T_2 < \dots$ de paramètre $\lambda > 0$. Chaque poisson est rouge avec probabilité p et gris avec probabilité $1 - p$. Si N_t^G (resp. N_t^R) est le nombre de poissons gris (resp. rouges) pêchés durant $[0, t]$, les processus N^G et N^R sont deux processus de Poisson indépendants, de paramètres $\lambda(1 - p)$ et λp .

2.4 La file d'attente $M/G/\infty$

Des clients arrivent aux instants $0 < T_1 < T_2 < \dots$ formant un processus de Poisson de paramètre λ . Le temps de service du client n est une variable aléatoire Y_n de loi η (sur $\mathbb{R}^+$). On suppose que la famille $(Y_n)_{n \geq 1}$ est i.i.d. et indépendante de $(T_n)_{n \geq 1}$. On considère le cas où il n'a pas d'attente: tout client est immédiatement servi. Ceci revient à supposer qu'il y a une infinité de serveurs.

On emploie la notation $M/G/\infty$ pour cette situation: le M indique que les arrivées sont poissonniennes (M est pour Markov, nous verrons plus tard pourquoi). Le G indique que la loi des services est arbitraire (G pour Général). Enfin ∞ indique qu'il y a une infinité de serveurs.

Théorème 2.12. Le nombre X_t de clients présents dans le système à l'instant t suit une loi de Poisson de paramètre $\lambda \int_0^\infty (y \wedge t) \eta(dy)$.

Preuve. On introduit le processus $\{(T_n, Y_n), n \geq 1\}$. C'est un processus ponctuel de Poisson sur $\mathbb{R}^+ \times \mathbb{R}^+$ d'intensité $\lambda m_+ \otimes \eta$, où m_+ est la mesure de Lebesgue sur $\mathbb{R}^+$. Remarquons ensuite que

$$X_t = \sum_{n=1}^{+\infty} \mathbf{1}_{\{T_n \leq t < T_n + Y_n\}} = \sum_{n=1}^{+\infty} \mathbf{1}_{\{(T_n, Y_n) \in I_t\}},$$

où $I_t = \{(s, y); 0 \leq s \leq t \leq s + y\}$. Autrement dit, en notant N la fonction de comptage du PPP, on a $X_t = N_{I_t}$. Par définition d'un processus ponctuel de Poisson, X_t a une loi de Poisson de paramètre $(\lambda m_+ \otimes \eta)(I_t)$. Or

$$\begin{aligned} (\lambda m_+ \otimes \eta)(I_t) &= \lambda \int_0^\infty \int_0^\infty \mathbf{1}_{I_t}(s, y) ds \eta(dy) \\ &= \lambda \int_0^\infty \left(\int_0^\infty \mathbf{1}_{0 < s < t < s+y} ds \right) \eta(dy) \\ &= \lambda \int_0^\infty \left(\int_{(t-y) \vee 0}^t ds \right) \eta(dy) \\ &= \lambda \int_0^\infty (y \wedge t) \eta(dy), \end{aligned}$$

car $t - (t - y) \vee 0 = y \wedge t$, séparer les cas $y \leq t$ et $y > t$. □

Remarque (Comportement en temps long). On déduit facilement de ce théorème que :

- (i) si le temps de service moyen $\mathbb{E}Y_1 = \int_0^\infty y \eta(dy)$ est fini, alors X_t converge en loi, quand $t \rightarrow \infty$, vers la loi de Poisson de paramètre $\lambda \int_0^\infty y \eta(dy)$. Le système est donc stable (n'explose pas);
- (ii) si le temps de service moyen $\mathbb{E}Y_1$ est infini, alors $X_t \rightarrow \infty$ en probabilité quand $t \rightarrow \infty$.

Processus régénératifs

On trouve dans la littérature de nombreuses définitions des processus régénératifs. Nous avons choisi celle qui est la plus adaptée aux théorèmes ergodiques. Dans la suite $\mathbf{T}$ désigne soit l'ensemble $\mathbb{N}$ des entiers soit l'ensemble $\mathbb{R}^+$. Nous supposons toujours que les processus $\{X_t, t \in \mathbb{R}^+\}$ à valeurs dans un espace topologique E , sont continus à droite.

Définition. On dit qu'un processus $(X_t)_{t \in \mathbf{T}}$, est régénératif s'il existe une suite croissante $\{\tau_n, n \in \mathbb{N}\}$ de variables aléatoires à valeurs dans $\mathbf{T}$ telle que, pour toute fonction mesurable $f : E \rightarrow \mathbb{R}^+$, les variables aléatoires

$$Z_n = \int_{\tau_n}^{\tau_{n+1}} f(X_s) ds \text{ si } \mathbf{T} = \mathbb{R}^+, \quad \text{ou} \quad Z_n = \sum_{k=\tau_n}^{\tau_{n+1}-1} f(X_k) \text{ si } \mathbf{T} = \mathbb{N},$$

sont i.i.d.

Très souvent $\tau_0 = 0$, mais ce n'est pas une nécessité. En prenant pour f la fonction identiquement égale à 1, on voit que les variables aléatoires $\{\tau_{n+1} - \tau_n, n \in \mathbb{N}\}$, sont positives, indépendantes et de même loi. On voit donc que τ_n est une marche aléatoire sur $\mathbb{R}$ à valeurs positives. On appelle parfois un tel processus un processus de renouvellement.

Exemple 3.1. Considérons une file d'attente où les délais d'arrivée $X_1, X_2, \dots$ entre les clients sont i.i.d. (de loi commune μ sur $\mathbb{R}^+$). Autrement dit, les temps d'arrivée des clients sont $T_1 = X_1$, $T_2 = X_1 + X_2$, $T_3 = X_1 + X_2 + X_3$, etc. Soit $N_t = \sum_{n \geq 1} \mathbf{1}_{\{T_n \leq t\}}$ le nombre de clients arrivés avant l'instant $t \geq 0$ (précisément, au plus tard au temps t). Alors les processus $X_t = t - T_{N_t}$ (temps écoulé, à l'instant t , depuis la dernière arrivée) et $Y_t = T_{N_t+1} - t$ (temps résiduel d'attente, à l'instant t , avant la prochaine arrivée), sont des processus régénératifs, et on peut choisir $\tau_n = T_n$.

En effet, considérons f mesurable positive sur $\mathbb{R}^+$, alors les v.a.

$$Z_n = \int_{T_n}^{T_{n+1}} f(X_s) ds = \int_{T_n}^{T_{n+1}} f(s - T_n) ds = \int_0^{T_{n+1} - T_n} f(u) du = \int_0^{X_{n+1}} f(u) du$$

sont bien entendu i.i.d., donc $(X_t)_{t \geq 0}$ est régénératif. On a utilisé le changement de variables $u =$

$s - T_n$.

De même, les v.a.

$$Z_n = \int_{T_n}^{T_{n+1}} f(Y_s) ds = \int_{T_n}^{T_{n+1}} f(T_{n+1} - s) ds = \int_0^{T_{n+1} - T_n} f(u) du = \int_0^{X_{n+1}} f(u) du$$

sont bien entendu i.i.d., donc $(Y_t)_{t \geq 0}$ est régénératif. On a utilisé le changement de variables $u = T_{n+1} - s$.

Remarque. On se gardera de penser que si $(X_t)_{t \geq 0}$ est un processus régénératif, alors les processus $E_n := (X_{\tau_n + s}, 0 \leq s \leq \tau_{n+1} - \tau_n)$, $n \in \mathbb{N}$, sont i.i.d. En effet, par exemple dans le cas discret où $\mathbf{T} = \mathbb{N}$, si $(Y_k)_{k \geq 0}$ est obtenu en permutant les valeurs du processus X de manière *arbitraire* sur les intervalles $\tau_n \leq k < \tau_{n+1}$ alors Y est aussi régénératif. Les processus $(E_n)_{n \geq 0}$ peuvent même être dépendants! On peut dans Y trier les valeurs sur chacun des intervalles (disons tous dans l'ordre croissant avec probabilité 1/2, ou bien tous dans l'ordre décroissant avec probabilité 1/2).

Nous utiliserons la version suivante de la loi des grands nombres.

Théorème 3.1. Soit $(Y_k)_{k \geq 1}$ des v.a. i.i.d. à valeurs positives. Presque sûrement,

$$\lim_{n \rightarrow +\infty} \frac{Y_1 + Y_2 + \dots + Y_n}{n} = \mathbf{E}(Y_1) \leq \infty.$$

Preuve. Lorsque $\mathbf{E}(Y_1) < +\infty$, c'est la loi des grands nombres classique. Si $\mathbf{E}(Y_1) = +\infty$, on remarque que pour tout $a > 0$, puisque $\min(Y_1, a)$ est d'espérance finie, p.s.,

$$\liminf_{n \rightarrow +\infty} \frac{\sum_{k=1}^n Y_k}{n} \geq \lim_{n \rightarrow +\infty} \frac{\sum_{k=1}^n \min(Y_k, a)}{n} = \mathbf{E}(\min(Y_1, a))$$

Ceci étant vrai pour tout a , par application du théorème de convergence monotone,

$$\liminf_{n \rightarrow +\infty} \frac{\sum_{k=1}^n Y_k}{n} \geq \lim_{a \rightarrow +\infty} \mathbf{E}(\min(Y_1, a)) = \infty \quad \text{p.s.} \quad \square$$

Théorème 3.2 (Théorème ergodique des processus régénératifs, $\mathbf{T} = \mathbb{R}^+$). Considérons un processus régénératif $(X_t)_{t \in \mathbb{R}^+}$ à valeurs dans E . Pour toutes fonctions mesurables $f, g : E \rightarrow \mathbb{R}^+$, presque sûrement,

$$\lim_{t \rightarrow +\infty} \frac{\int_0^t f(X_s) ds}{\int_0^t g(X_s) ds} = \frac{\mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s) ds]}{\mathbf{E}[\int_{\tau_0}^{\tau_1} g(X_s) ds]},$$

si $0 < \mathbf{E}[\int_{\tau_0}^{\tau_1} g(X_s) ds] < +\infty$. En particulier, si $0 < \mathbf{E}[\tau_1 - \tau_0] < +\infty$,

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t f(X_s) ds = \frac{\mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s) ds]}{\mathbf{E}[\tau_1 - \tau_0]} = \int_E f(x) \pi(dx),$$

où π est la probabilité sur E définie par (pour $A \in \mathcal{E}$)

$$\pi(A) = \frac{1}{\mathbf{E}[\tau_1 - \tau_0]} \mathbf{E}\left[\int_{\tau_0}^{\tau_1} \mathbf{1}_{\{X_s \in A\}} ds\right].$$

Preuve. Etape 1. Posons $N_t = \sum_{k=1}^{+\infty} \mathbf{1}_{\{\tau_k \leq t\}}$. Alors N_t tend p.s. vers l'infini quand $t \rightarrow \infty$ (car N_t est croissant et car $N_{\tau_k} = k$). De plus, pour tout $t \geq \tau_0$, par définition de N_t ,

$$\tau_{N_t} \leq t \leq \tau_{N_t+1}.$$

On a donc, puisque f est à valeurs positives,

$$\int_0^{\tau_{N_t}} f(X_s) ds \leq \int_0^t f(X_s) ds \leq \int_0^{\tau_{N_t+1}} f(X_s) ds.$$

Etape 2. On écrit

$$\frac{1}{N_t} \int_0^{\tau_{N_t}} f(X_s) ds = \frac{1}{N_t} \int_0^{\tau_0} f(X_s) ds + \frac{1}{N_t} \sum_{k=0}^{N_t-1} \int_{\tau_k}^{\tau_{k+1}} f(X_s) ds.$$

Par hypothèse, les v.a. $Z_k = \int_{\tau_{k-1}}^{\tau_k} f(X_s) ds$ sont i.i.d. (et positives), donc par la LGN, $\frac{1}{n} \sum_{k=0}^{n-1} Z_k$ converge p.s. vers $\mathbf{E}(Z_1) = \mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s) ds]$. Quand $t \rightarrow +\infty$, puisque $N_t \rightarrow +\infty$ p.s., on en déduit que $\frac{1}{N_t} \sum_{k=0}^{N_t-1} Z_k$ converge p.s. vers la même limite $\mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s) ds]$. Comme enfin $\frac{1}{N_t} \int_0^{\tau_0} f(X_s) ds$ tend vers 0 p.s. (puisque $N_t \rightarrow \infty$), on conclut que p.s.,

$$\lim_{t \rightarrow \infty} \frac{1}{N_t} \int_0^{\tau_{N_t}} f(X_s) ds = \mathbf{E} \left[\int_{\tau_0}^{\tau_1} f(X_s) ds \right].$$

Etape 3. On montre de même (utiliser aussi que $(N_t + 1)/N_t \rightarrow 1$) que p.s., quand $t \rightarrow \infty$

$$\lim_{t \rightarrow \infty} \frac{1}{N_t} \int_0^{\tau_{N_t+1}} f(X_s) ds = \lim_{t \rightarrow \infty} \frac{N_t + 1}{N_t} \frac{1}{N_t + 1} \int_0^{\tau_{N_t+1}} f(X_s) ds = \mathbf{E} \left[\int_{\tau_0}^{\tau_1} f(X_s) ds \right].$$

Etape 4. Les étapes 1,2,3 montrent que pour f positive, on a p.s.

$$\lim_{t \rightarrow \infty} \frac{1}{N_t} \int_0^t f(X_s) ds = \mathbf{E} \left[\int_{\tau_0}^{\tau_1} f(X_s) ds \right].$$

Supposons maintenant qu'on a 2 fonctions positives f, g et que $0 < \mathbf{E}[\int_{\tau_0}^{\tau_1} g(X_s) ds] < \infty$. On écrit alors

$$\lim_{t \rightarrow \infty} \frac{\int_0^t f(X_s) ds}{\int_0^t g(X_s) ds} = \lim_{t \rightarrow \infty} \frac{N_t^{-1} \int_0^t f(X_s) ds}{N_t^{-1} \int_0^t g(X_s) ds} = \frac{\mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s) ds]}{\mathbf{E}[\int_{\tau_0}^{\tau_1} g(X_s) ds]}.$$

Finalement, si on peut appliquer ce résultat avec $g = 1$, i.e. si $0 < \mathbf{E}[\tau_1 - \tau_0] < \infty$, alors on trouve que pour toute f mesurable positive,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(X_s) ds = \frac{\mathbf{E}[\int_{\tau_0}^{\tau_1} f(X_s) ds]}{\mathbf{E}[\tau_1 - \tau_0]}.$$

Le membre de droite n'est autre que $\int_E f d\pi$. □

Exemple 3.2 (Processus de renouvellement alterné). Modélisons le comportement d'une machine qui est successivement en état de marche et en panne comme suit. La machine tombe en panne, après la n -ième réparation, au bout d'un temps X_{n+1} . Se fait réparer, après la n -ième panne, pendant une durée Y_n . Autrement dit, la machine marche du temps $T_0 = 0$ au temps $T_1 = X_1$, elle est en réparation du temps T_1 au temps $T_2 = X_1 + Y_1$, puis elle marche du temps T_2 au temps $T_3 = T_2 + X_2$, puis elle est en réparation du temps T_3 au temps $T_4 = T_3 + Y_2$, etc.

Soit $A_t = 1$ si la machine marche à l'instant t , $A_t = 0$ sinon. On suppose que les v.a. X_i sont i.i.d. (positives), indépendantes des Y_i (positives) aussi i.i.d. On suppose aussi que $0 < \mathbf{E}[X_1] < \infty$ et que $0 < \mathbf{E}[Y_1] < \infty$.

Le processus $(A_t)_{t \geq 0}$ est régénératif et on peut prendre $\tau_n = T_{2n}$: en effet, pour f mesurable

positive, les v.a.

$$\begin{aligned} Z_n &= \int_{T_{2n}}^{T_{2n+2}} f(A_s) ds \\ &= \int_{T_{2n}}^{T_{2n}+X_{n+1}} f(1) ds + \int_{T_{2n}+X_{n+1}}^{T_{2n}+X_{n+1}+Y_{n+1}} f(0) ds \\ &= f(1)X_{n+1} + f(0)Y_{n+1} \end{aligned}$$

sont bien sûr i.i.d. Donc on déduit du théorème que p.s.,

$$\lim_{t \rightarrow \infty} \frac{A_t}{t} = \frac{\mathbf{E}[\int_{\tau_0}^{\tau_1} A_s ds]}{\mathbf{E}[\tau_1 - \tau_0]} = \frac{\mathbf{E}[\int_0^{T_2} A_s ds]}{\mathbf{E}[T_2]} = \frac{\mathbf{E}[X_1]}{\mathbf{E}[X_1 + Y_1]}.$$

Autrement dit, la proportion du temps (sur un intervalle de temps infini) pendant laquelle la machine marche vaut $\mathbf{E}[X_1]/\mathbf{E}[X_1 + Y_1]$, ce qui est intuitivement clair.

On montre de la même façon la version suivante du théorème en temps discret:

Théorème 3.3 (Théorème ergodique des processus régénératifs, $T = \mathbb{N}$). Considérons un processus régénératif $(X_n)_{n \in \mathbb{N}}$ à valeurs dans E . Pour toutes fonctions mesurables $f, g : E \rightarrow \mathbb{R}^+$, presque sûrement,

$$\lim_{n \rightarrow +\infty} \frac{\sum_{k=0}^n f(X_k)}{\sum_{k=0}^n g(X_k)} = \frac{\mathbf{E}[\sum_{k=\tau_0}^{\tau_1-1} f(X_k)]}{\mathbf{E}[\sum_{k=\tau_0}^{\tau_1-1} g(X_k)]},$$

si $0 < \mathbf{E}[\sum_{k=\tau_0}^{\tau_1-1} g(X_k)] < +\infty$. Lorsque $0 < \mathbf{E}[\tau_1 - \tau_0] < +\infty$,

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^n f(X_k) = \frac{\mathbf{E}[\sum_{k=\tau_0}^{\tau_1-1} f(X_k)]}{\mathbf{E}[\tau_1 - \tau_0]} = \int_E f(x) \pi(dx),$$

où π est la probabilité sur E définie par (pour $A \in \mathcal{E}$)

$$\pi(A) = \frac{1}{\mathbf{E}[\tau_1 - \tau_0]} \mathbf{E} \left[\sum_{k=\tau_0}^{\tau_1-1} \mathbf{1}_{\{X_k \in A\}} \right].$$

Chaînes de Markov

Ce chapitre est une introduction générale aux chaînes de Markov. Dans tout le chapitre, E est un ensemble fini ou dénombrable, et $\mathcal{E}$ est la tribu de toutes les parties de E . L'objectif est de mettre en place les objets en temps discret en préparation du chapitre suivant, qui concerne les processus Markovien de sauts, qui sont des "chaînes de Markov à temps continu".

4.1 Matrices de transition

Définition. Une matrice (ou noyau, ou probabilité) de transition sur E est une famille $\{P(i, j), i, j \in E\}$ de réels telle que

- (i) $P(i, j) \geq 0$, pour tout $i, j \in E$;
- (ii) pour tout $i \in E$, $\sum_{j \in E} P(i, j) = 1$.

Commençons par quelques propriétés et notations. Soient P, Q deux matrices de transitions sur E , soit f une fonction de E dans $\mathbb{R}$ (on considère alors $f = (f(i))_{i \in E}$ comme un vecteur colonne) et soit μ une probabilité sur E (on considère alors $\mu = (\mu(i))_{i \in E}$ comme un vecteur ligne), alors:

- PQ est la matrice de transition définie par $PQ(i, j) = \sum_{k \in E} P(i, k)Q(k, j)$ (vérifier que c'est une matrice de transition).
- P^n est la matrice de transition définie par $P^0(i, j) = \mathbf{1}_{\{i=j\}}$ et $P^{n+1} = P^n P$.
- Pf est la fonction de E dans $\mathbb{R}$ définie par $Pf(i) = \sum_{j \in E} P(i, j)f(j)$.
- μP est la probabilité sur E définie par $\mu P(j) = \sum_{i \in E} \mu(i)P(i, j)$ (vérifier que c'est une probabilité).
- μf est le nombre $\sum_{i \in E} \mu(i)f(i)$. C'est $\mathbf{E}[f(X)]$, si X est une v.a. à valeurs dans E de loi μ .

Bref, tous les produits *matriciels* usuels sont utilisables dans ce contexte, en identifiant les fonctions à des vecteurs colonnes et les mesures à des vecteurs lignes.

Exercice 4.1. Vérifier que si P est une matrice de transition, f une fonction de E dans $\mathbb{R}$ et μ une mesure de probabilité sur E alors $\mu(Pf) = (\mu P)f$.

4.2 Chaîne de Markov

Définition (Chaîne de Markov). Soit P une matrice de transition sur E . Un processus $(X_n)_{n \in \mathbb{N}}$ à valeurs dans E est une **chaîne de Markov** de transition P si pour tout $n \geq 0$, tous $i_0, \dots, i_n \in E$, tout $j \in E$,

$$\mathbf{P}(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = P(i_n, j).$$

La loi de X_0 s'appelle la loi initiale de la chaîne.

Exercice 4.2. Montrer qu'on a alors $\mathbf{P}(X_{n+1} = j \mid X_n = i) = P(i, j)$ pour tout $n \geq 0$, tout $i, j \in E$.

Une chaîne de Markov vérifie donc deux propriétés importantes :

- (i) *propriété de Markov*: X_{n+1} ne dépend de tout le passé jusqu'à l'instant n qu'à travers la valeur de X_n ;
- (ii) *homogénéité*: la loi de X_{n+1} sachant X_n ne dépend pas du temps n .

Pour montrer qu'un processus est une chaîne de Markov, on utilise très souvent le lemme suivant.

Lemme 4.1. Soit E un ensemble dénombrable, $(\varepsilon_n)_{n \geq 1}$ une suite de v.a. i.i.d. à valeurs dans un ensemble F . Soit $\phi : E \times F \rightarrow E$ une application mesurable. Pour tout X_0 indépendant de $(\varepsilon_n)_{n \geq 1}$, la suite définie par récurrence par

$$X_{n+1} = \phi(X_n, \varepsilon_{n+1})$$

est une chaîne de Markov à valeurs dans E de transition $P(i, j) = \mathbf{P}(\phi(i, \varepsilon_1) = j)$.

Preuve. Ceci résulte du fait que

$$\begin{aligned} \mathbf{P}(X_{n+1} = j \mid X_0 = i_0, \dots, X_n = i_n) &= \mathbf{P}(\phi(X_n, \varepsilon_{n+1}) = j \mid X_0 = i_0, \dots, X_n = i_n) \\ &= \mathbf{P}(\phi(i_n, \varepsilon_{n+1}) = j \mid X_0 = i_0, \dots, X_n = i_n) \\ &= \mathbf{P}(\phi(i_n, \varepsilon_{n+1}) = j) \\ &= \mathbf{P}(\phi(i_n, \varepsilon_1) = j) = P(i_n, j). \end{aligned}$$

On a utilisé que ε_{n+1} est indépendant de $X_0, \dots, X_n$, puis que ε_{n+1} a la même loi que ε_1 . □

Remarque. Soit P une matrice de transition sur E et ν une loi sur E . On peut construire une chaîne de Markov de transition P et de loi initiale ν ainsi (on suppose $E = \{i_k, k \geq 1\}$) : on considère X_0 de loi ν , indépendant d'une famille $(U_n)_{n \geq 1}$ i.i.d. de v.a. uniformes sur $[0, 1]$, et on construit X_n par récurrence en posant $X_{n+1} = \phi(X_n, U_{n+1})$, où $\phi : E \times [0, 1] \rightarrow E$ est définie par :

$$\phi(i, u) = \sum_{k=1}^{\infty} i_k \mathbf{1}_{\{u \in J_i(k)\}}$$

avec $J_i(k) = [P(i, i_1) + \dots + P(i, i_{k-1}), P(i, i_1) + \dots + P(i, i_k)]$ sont des intervalles disjoints de longueurs $|J_i(k)| = P(i, i_k)$. Par le lemme, c'est une chaîne de Markov, et on vérifie aisément que sa matrice de transition est donnée par $\mathbf{P}(\phi(i, U_1) = j) = P(i, j)$.

Exemple 4.1. Rappelons qu'une marche aléatoire de pas de loi μ à valeurs dans $\mathbb{Z}^d$ est définie par $X_0 = 0$ et, pour $n \geq 1$, $X_n = Y_1 + \dots + Y_n$, où les Y_k sont i.i.d. de loi μ . En appliquant le lemme avec $E = F = \mathbb{Z}^d$, avec $\varepsilon_n = Y_n$, et avec $\phi(x, y) = x + y$ (on a bien $X_{n+1} = X_n + Y_{n+1} = \phi(X_n, Y_{n+1})$), on trouve que $(X_n)_{n \geq 0}$ est une chaîne de Markov à valeurs dans $\mathbb{Z}^d$ de transition $P(i, j) = \mathbf{P}(\phi(i, Y_1) = j) = \mu(\{j - i\})$ (c'est logique : pour passer de i à j , il faut sauter de $j - i$).

Exemple 4.2. Considérons une famille de v.a. i.i.d. $(Y_n)_{n \geq 1}$ de loi $(\delta_{-1} + \delta_1)/2$ et posons $X_0 = 0$ puis $X_n = Y_1 + \dots + Y_n$ pour $n \geq 1$. Alors $(X_n)_{n \geq 0}$ est la marche aléatoire simple symétrique sur $\mathbb{Z}$, c'est donc une chaîne de Markov. Par contre, $S_n = \sup\{X_k, k = 0, \dots, n\}$ n'est pas une chaîne de Markov : on peut se convaincre, par exemple, que

$$\mathbf{P}(S_4 = 2 \mid S_0 = 0, S_1 = 1, S_2 = 1, S_3 = 1) \neq \mathbf{P}(S_4 = 2 \mid S_0 = 0, S_1 = 0, S_2 = 0, S_3 = 1).$$

La probabilité de gauche vaut $1/4$, celle de droite $1/2$.

Remarque. Attention, il est faux en général, pour $(X_n)_{n \geq 0}$ une chaîne de Markov, que $\mathbf{P}(X_2 \in A_2 \mid X_0 \in A_0, X_1 \in A_1) = \mathbf{P}(X_2 \in A_2 \mid X_1 \in A_1)$ si les ensembles A_i ne sont pas des singletons. En effet, si $E = \{1, 2, 3, 4, 5, 6\}$, si $P(1, 2) = P(2, 3) = P(3, 3) = 1$ et $P(4, 5) = P(5, 6) = P(6, 6) = 1$, et si on a $A_0 = \{1\}$, $A_1 = \{2, 5\}$ et $A_2 = \{3\}$, alors $\mathbf{P}(X_2 \in A_2 \mid X_0 \in A_0, X_1 \in A_1) = 1$. Mais avec $A'_0 = \{4\}$, $A_1 = \{2, 5\}$ et $A_2 = \{3\}$, $\mathbf{P}(X_2 \in A_2 \mid X_0 \in A'_0, X_1 \in A_1) = 0$.

Notation. On considèrera souvent plusieurs lois initiales pour la même matrice de transition. Il nous faudra donc souvent préciser avec quelle loi initiale on travaille. On notera en indice cette loi initiale dans les espérances et les probabilités, $\mathbf{E}_\nu$, $\mathbf{P}_\nu$. Par exemple, si on travaille avec une chaîne de Markov $(X_n)_{n \geq 0}$ de transition P , on note $\mathbf{E}_\nu(X_k)$ (resp. $\mathbf{P}_\nu(X_k \in A)$) l'espérance de X_k (resp. la probabilité que $X_k \in A$) si $(X_n)_{n \geq 0}$ est la chaîne de transition P et de loi initiale ν .

Dans le cas où $\nu = \delta_x$ pour un $x \in E$ (i.e. quand $X_0 = x$ p.s.), on simplifie la notation et on écrit $\mathbf{P}_x, \mathbf{E}_x$ pour $\mathbf{P}_{\delta_x}, \mathbf{E}_{\delta_x}$.

Lemme 4.2. Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de transition P . Pour toute probabilité ν sur E ,

$$\mathbf{P}_\nu(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \nu(x_0)P(x_0, x_1) \cdots P(x_{n-1}, x_n).$$

En particulier (avec $\nu = \delta_x$),

$$\mathbf{P}_x(X_1 = x_1, \dots, X_n = x_n) = P(x, x_1) \cdots P(x_{n-1}, x_n).$$

Preuve. On procède par récurrence sur $n \in \mathbb{N}$: si $n = 0$, c'est évident, puisque $\mathbf{P}_\nu(X_0 = x_0) = \nu(x_0)$ par définition de $\mathbf{P}_\nu$. Si la formule est vraie avec $n \geq 0$, alors on écrit

$$\begin{aligned} \mathbf{P}_\nu(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, X_{n+1} = x_{n+1}) \\ = \mathbf{P}_\nu(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) \times \mathbf{P}_\nu(X_{n+1} = x_{n+1} \mid X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) \\ = \nu(x_0)P(x_0, x_1) \cdots P(x_{n-1}, x_n) \times P(x_n, x_{n+1}). \end{aligned}$$

□

Corollaire 4.3. La loi d'une chaîne de Markov est entièrement caractérisée par sa matrice de transition et sa loi initiale.

Proposition 4.4. Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov de transition P et de loi initial ν , et soit $n \geq 0$ fixé.

- (i) La loi de X_n est νP^n , i.e. $\mathbf{P}_\nu(X_n = y) = \nu P^n(y)$ pour tout $y \in E$. En particulier, $\mathbf{P}_x(X_n = y) = P^n(x, y)$ et $P^n(x, y)$ est la probabilité que la chaîne passe de x à y en n pas.
- (ii) Pour $f : E \rightarrow \mathbb{R}^+$, on a $\mathbf{E}_\nu(f(X_n)) = \nu P^n f$. En particulier, $\mathbf{E}_x(f(X_n)) = P^n f(x)$.

Preuve. (ii) découle de (i). Pour (i), fixons $y \in E$, et écrivons, comme E est dénombrable,

$$\begin{aligned} \mathbf{P}_\nu(X_n = y) &= \sum_{x_0, \dots, x_{n-1} \in E} \mathbf{P}_\nu(X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = y) \\ &= \sum_{x_0, \dots, x_{n-1} \in E} \nu(x_0, x_1) P(x_1, x_2) \cdots P(x_{n-2}, x_{n-1}) P(x_{n-1}, y). \end{aligned}$$

En utilisant les produits matriciels, ceci donne $\nu P^n(y)$ (rappelons que ν est un vecteur ligne et P une matrice, donc νP^n un vecteur ligne). Avec le choix particulier $\nu = \delta_x$, on trouve

$$\mathbf{P}_x(X_n = y) = (\delta_x P^n)(y) = \sum_{z \in E} \delta_x(z) P^n(z, y) = P^n(x, y). \quad \square$$

4.3 Propriété de Markov

Soit $(X_n)_{n \geq 0}$ une chaîne de Markov sur E . On note $\mathcal{F}_n$ la tribu engendrée par $(X_0, X_1, \dots, X_n)$.

Théorème 4.5 (Propriété de Markov simple). Pour toute fonction mesurable $f : E^\mathbb{N} \rightarrow \mathbb{R}$, par exemple bornée ou positive, pour toute probabilité ν , presque sûrement,

$$\mathbf{E}_\nu[f(X_n, X_{n+1}, \dots) \mid \mathcal{F}_n] = \mathbf{E}_{X_n}[f(X_0, X_1, \dots)].$$

Ceci signifie que pour tout $x_0, x_1, \dots, x_n \in E$, on a

$$\mathbf{E}_\nu[f(X_n, X_{n+1}, \dots) \mid X_0 = x_0, \dots, X_n = x_n] = \mathbf{E}_{x_n}[f(X_0, X_1, \dots)].$$

Preuve. Il suffit (voir Section A.3) de vérifier ce résultat lorsque f est une fonction indicatrice $f = \mathbf{1}_A$ où A est de la forme

$$A = \{a_0\} \times \{a_1\} \times \cdots \times \{a_d\} \times E \times E \times \cdots.$$

On a alors $f(X_n, X_{n+1}, \dots) = \mathbf{1}_{\{X_n = a_0, X_{n+1} = a_1, \dots, X_{n+d} = a_d\}}$. L'objectif est donc de démontrer que pour tous $a_0, \dots, a_d \in E$ et pour tous $x_0, \dots, x_n \in E$, on a $I = J$, où

$$I = \mathbf{P}_\nu(X_n = a_0, X_{n+1} = a_1, \dots, X_{n+d} = a_d \mid X_0 = x_0, X_1 = x_1, \dots, X_n = x_n)$$

et

$$J = \mathbf{P}_{x_n}(X_0 = a_0, X_1 = a_1, \dots, X_d = a_d).$$

Si $a_0 \neq x_n$, c'est évident, puisque $I = 0 = J$. Si maintenant $a_0 = x_n$, on a d'une part

$$J = P(a_0, a_1) P(a_1, a_2) \cdots P(a_{d-1}, a_d)$$

et d'autre part

$$\begin{aligned}
 I &= \frac{\mathbf{P}_v(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, X_{n+1} = a_1, \dots, X_{n+d} = a_d)}{\mathbf{P}_v(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n)} \\
 &= \frac{v(x_0)P(x_0, x_1) \cdots P(x_{n-1}, x_n)P(x_n, a_1) \cdots P(a_{d-1}, a_d)}{v(x_0)P(x_0, x_1) \cdots P(x_{n-1}, x_n)} \\
 &= P(x_n, a_1) \cdots P(a_{d-1}, a_d).
 \end{aligned}$$

Comme $x_n = a_0$, on a bien $I = J$. □

Rappelons qu'un temps d'arrêt du processus $(X_n)_{n \geq 0}$ est une v.a. à valeurs dans $\mathbb{N} \cup \{\infty\}$ telle que pour tout $n \in \mathbb{N}$, l'événement $\{T = n\}$ est dans $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$. La tribu $\mathcal{F}_T$ est formée des événements $A \in \mathcal{F}_\infty$ tels que pour tout $n \in \mathbb{N}$, $A \cap \{T = n\} \in \mathcal{F}_n$. En particulier $T, \mathbf{1}_{\{T < \infty\}}X_T$ et $X_{T \wedge n}$ sont des variables aléatoires $\mathcal{F}_T$ -mesurables.

Lemme 4.6. Soit T un temps d'arrêt. Pour toute variable aléatoire Z bornée (ou positive), pour tout $n \geq 0$,

$$\mathbf{1}_{\{T=n\}}\mathbf{E}(Z \mid \mathcal{F}_T) = \mathbf{1}_{\{T=n\}}\mathbf{E}(Z \mid \mathcal{F}_n).$$

Preuve. Montrons d'abord que $\{T = n\} \in \mathcal{F}_T$. Pour cela il faut vérifier que pour tout $k \geq 0$, on a $\{T = n\} \cap \{T = k\} \in \mathcal{F}_k$. C'est vrai si $k \neq n$ (car alors $\{T = n\} \cap \{T = k\} = \emptyset \in \mathcal{F}_k$) et si $k = n$ (car alors $\{T = n\} \cap \{T = k\} = \{T = k\} \in \mathcal{F}_k$ puisque T est un temps d'arrêt).

Du coup, $\mathbf{1}_{\{T=n\}}\mathbf{E}[Z \mid \mathcal{F}_T] = \mathbf{E}[\mathbf{1}_{\{T=n\}}Z \mid \mathcal{F}_T]$. Il faut donc montrer que

$$\mathbf{E}[\mathbf{1}_{\{T=n\}}Z \mid \mathcal{F}_T] = \mathbf{1}_{\{T=n\}}\mathbf{E}[Z \mid \mathcal{F}_n],$$

ce qui signifie que

- (i) $\mathbf{1}_{\{T=n\}}\mathbf{E}[Z \mid \mathcal{F}_n]$ est $\mathcal{F}_T$ -mesurable et
- (ii) pour tout $A \in \mathcal{F}_T$, $\mathbf{E}[\mathbf{1}_A \mathbf{1}_{\{T=n\}}Z] = \mathbf{E}[\mathbf{1}_A \mathbf{1}_{\{T=n\}}\mathbf{E}(Z \mid \mathcal{F}_n)]$.

Pour (i), il faut vérifier que pour tout $B \in \mathcal{B}(\mathbb{R})$, $\{\mathbf{1}_{\{T=n\}}\mathbf{E}[Z \mid \mathcal{F}_n] \in B\} \in \mathcal{F}_T$, soit encore que pour tout $k \geq 0$, $\{\mathbf{1}_{\{T=n\}}\mathbf{E}[Z \mid \mathcal{F}_n] \in B\} \cap \{T = k\} \in \mathcal{F}_k$: c'est juste si $k = n$, puisque $\{T = n\} \in \mathcal{F}_n$ et puisque $\mathbf{E}[Z \mid \mathcal{F}_n]$ est $\mathcal{F}_n$ -mesurable ; c'est aussi juste si $k \neq n$, puisqu'alors $\{\mathbf{1}_{\{T=n\}}\mathbf{E}[Z \mid \mathcal{F}_n] \in B\} \cap \{T = k\} = \{0 \in B\} \cap \{T = k\} \in \mathcal{F}_k$ (noter que $\{0 \in B\}$ vaut soit $\emptyset$ soit Ω).

Pour montrer (ii), on utilise le fait que $\mathbf{1}_A \mathbf{1}_{\{T=n\}}$ est $\mathcal{F}_n$ mesurable (car $A \cap \{T = n\} \in \mathcal{F}_n$ car $A \in \mathcal{F}_T$ par hypothèse), d'où on déduit que

$$\mathbf{E}[\mathbf{1}_A \mathbf{1}_{\{T=n\}}\mathbf{E}(Z \mid \mathcal{F}_n)] = \mathbf{E}[\mathbf{E}(\mathbf{1}_A \mathbf{1}_{\{T=n\}}Z \mid \mathcal{F}_n)] = \mathbf{E}[\mathbf{1}_A \mathbf{1}_{\{T=n\}}Z].$$
 □

On déduit la propriété de Markov forte des deux résultats précédents.

Théorème 4.7 (Propriété de Markov forte). Soit T un temps d'arrêt de la chaîne de Markov $(X_n)_{n \geq 0}$. Pour toute fonction mesurable $f : E^\mathbb{N} \rightarrow \mathbb{R}$ bornée ou positive, pour toute probabilité v sur E , p.s.

$$\mathbf{1}_{\{T < \infty\}}\mathbf{E}_v[f(X_T, X_{T+1}, \dots) \mid \mathcal{F}_T] = \mathbf{1}_{\{T < \infty\}}\mathbf{E}_{X_T}[f(X_0, X_1, \dots)].$$

Preuve. Par le lemme, on a, pour tout $n \geq 0$,

$$\begin{aligned} \mathbf{1}_{\{T=n\}} \mathbf{E}_v[f(X_T, X_{T+1}, \dots) \mid \mathcal{F}_T] &= \mathbf{1}_{\{T=n\}} \mathbf{E}_v[f(X_T, X_{T+1}, \dots) \mid \mathcal{F}_n] \\ &= \mathbf{1}_{\{T=n\}} \mathbf{E}_v[f(X_n, X_{n+1}, \dots) \mid \mathcal{F}_n] \\ &= \mathbf{1}_{\{T=n\}} \mathbf{E}_{X_n}[f(X_0, X_1, \dots)] \\ &= \mathbf{1}_{\{T=n\}} \mathbf{E}_{X_T}[f(X_0, X_1, \dots)], \end{aligned}$$

où on a utilisé Markov simple à l'avant-dernière ligne. En sommant sur $n \in \mathbb{N}$ et en utilisant le fait que $\sum_{n \geq 0} \mathbf{1}_{\{T=n\}} = \mathbf{1}_{\{T < \infty\}}$, on trouve

$$\mathbf{1}_{\{T < \infty\}} \mathbf{E}_v[f(X_T, X_{T+1}, \dots) \mid \mathcal{F}_T] = \mathbf{1}_{\{T < \infty\}} \mathbf{E}_{X_T}[f(X_0, X_1, \dots)]. \quad \square$$

Notons la conséquence suivante de la propriété de Markov forte.

Corollaire 4.8. Soit T un temps d'arrêt de la chaîne de Markov $(X_n)_{n \geq 0}$, fini presque sûrement et tel que $X_T = x$ p.s. (pour un certain $x \in E$). Alors, $(X_{T+n})_{n \geq 0}$ est une chaîne de Markov de transition P (la même que celle de $(X_n)_{n \geq 0}$), de loi initiale δ_x , qui est de plus indépendante de $\mathcal{F}_T$.

Preuve. Il suffit de montrer que, pour tout $A \in \mathcal{F}_T$ et tous $x_1, \dots, x_k \in E$, on a $I = J$, où

$$I = \mathbf{P}_v(A, X_T = x, X_{T+1} = x_1, \dots, X_{T+k} = x_k) \quad \text{et} \quad J = \mathbf{P}_v(A)P(x, x_1) \cdots P(x_{k-1}, x_k).$$

Mais, comme $A \in \mathcal{F}_T$,

$$\begin{aligned} I &= \mathbf{E}_v[\mathbf{1}_A \mathbf{E}[\mathbf{1}_{\{X_T=x, X_{T+1}=x_1, \dots, X_{T+k}=x_k\}} \mid \mathcal{F}_T]] \\ &= \mathbf{E}_v[\mathbf{1}_A \mathbf{E}_{X_T}[\mathbf{1}_{\{X_0=x, X_1=x_1, \dots, X_k=x_k\}}]] \end{aligned}$$

par Markov forte. Puis, comme $X_T = x$,

$$\begin{aligned} I &= \mathbf{E}_v[\mathbf{1}_A \mathbf{E}_x[\mathbf{1}_{\{X_0=x, X_1=x_1, \dots, X_k=x_k\}}]] \\ &= \mathbf{E}_v[\mathbf{1}_A P(x, x_1) \cdots P(x_{k-1}, x_k)], \end{aligned}$$

qui n'est autre que J . □

4.4 Propriétés de récurrence

Nous allons étudier le comportement des trajectoires d'une chaîne de Markov X_n lorsque $n \rightarrow +\infty$. On va distinguer deux types de comportement suivant que la chaîne "tend vers l'infini" ou non. Introduisons les notations suivantes : pour $i \in E$,

$$\tau_i = \inf\{k > 0; X_k = i\}.$$

Puis $\tau_i^{(0)} = 0$ et, pour tout entier $\ell \geq 0$,

$$\tau_i^{(\ell+1)} = \inf\{k > \tau_i^{(\ell)}; X_k = i\}.$$

On voit que $\tau_i^{(1)} = \tau_i$ est le temps d'atteinte de l'état i (en excluant si nécessaire le temps $n = 0$), puis $\tau_i^{(\ell)}$ est l'instant de ℓ -ième passage en i (en excluant si nécessaire le temps $n = 0$). Ce sont des temps d'arrêt.

Pour tout $i, j \in E$, on introduit

$$G(i, j) = \sum_{n=0}^{+\infty} P^n(i, j) = \sum_{n=0}^{+\infty} \mathbf{E}_i(\mathbf{1}_{\{X_n=j\}}) = \mathbf{E}_i \left[\sum_{n=0}^{+\infty} \mathbf{1}_{\{X_n=j\}} \right].$$

C'est l'espérance du nombre de passages par j en partant de i .

Définition. On dit qu'un état i est **récurrent** si $\mathbf{P}_i(\tau_i < +\infty) = 1$.

Théorème 4.9. Les assertions suivantes sont équivalentes:

1. i est récurrent.
2. $\sum_{n=0}^{+\infty} \mathbf{1}_{\{X_n=i\}} = +\infty$, $\mathbf{P}_i$ -p.s.
3. $G(i, i) = +\infty$.

On a donc, pour un état i donné, l'équivalence : 1. partant de i , on est sûr de revenir en i ; 2. partant de i , on est sûr de revenir une infinité de fois en i ; 3. le nombre moyen de passages en i , partant de i , est infini.

Remarque. L'équivalence entre 1 et 2 est intuitivement claire, et bien sûr 2 implique 3. Par contre, il est assez surprenant que le nombre de passages par i soit infini p.s. dès que son espérance est infinie (c'est ce que dit 3 implique 2).

Preuve. Etape 1. Soit $N = \sum_{n=0}^{+\infty} \mathbf{1}_{\{X_n=i\}}$ le nombre de passages par i . On va montrer ici que, sous $\mathbf{P}_i$, N a une loi géométrique, c'est-à-dire que pour tout $n \geq 0$, $\mathbf{P}_i(N > n) = \alpha^n$, où $\alpha = \mathbf{P}_i(\tau_i < +\infty)$.

C'est clair si $n = 0$ (car comme on part de i , on a $N \geq 1$ $\mathbf{P}_i$ -p.s.). Il suffit donc de montrer que pour tout $n \geq 0$, $\mathbf{P}_i(N > n+1) = \alpha \mathbf{P}_i(N > n)$. On écrit donc

$$\begin{aligned} \mathbf{P}_i(N > n+1) &= \mathbf{P}_i(\tau_i^{(n+1)} < \infty) \\ &= \mathbf{E}_i \left[\mathbf{P}_i(\tau_i^{(n+1)} < \infty \mid \mathcal{F}_{\tau_i^{(n)}}) \right] \\ &= \mathbf{E}_i \left[\mathbf{1}_{\{\tau_i^{(n)} < \infty\}} \mathbf{P}_i(\tau_i^{(n+1)} - \tau_i^{(n)} < \infty \mid \mathcal{F}_{\tau_i^{(n)}}) \right]. \end{aligned} \quad (\star)$$

La dernière égalité vient du fait que $\tau_i^{(n+1)} < \infty$ ssi $\tau_i^{(n)} < \infty$ et $\tau_i^{(n+1)} - \tau_i^{(n)} < \infty$ et du fait que $\tau_i^{(n)}$ est $\mathcal{F}_{\tau_i^{(n)}}$ -mesurable.

Par Markov forte, $(\tilde{X}_k = X_{\tau_i^{(n)}+k})_{k \geq 0}$ est une chaîne de Markov de même transition que $(X_k)_{k \geq 0}$, qui part de i et qui est indépendante de $\mathcal{F}_{\tau_i^{(n)}}$. Donc, comme $\tilde{\tau}_i = \inf\{k \geq 1 : \tilde{X}_k = i\}$ est égal à $\tau_i^{(n+1)} - \tau_i^{(n)}$, on conclut que

$$\mathbf{1}_{\{\tau_i^{(n)} < \infty\}} \mathbf{P}_i(\tau_i^{(n+1)} - \tau_i^{(n)} < \infty \mid \mathcal{F}_{\tau_i^{(n)}}) = \mathbf{1}_{\{\tau_i^{(n)} < \infty\}} \mathbf{P}_i(\tau_i < \infty). \quad (\star\star)$$

Ainsi, en revenant à $(\star)$,

$$\mathbf{P}_i(N > n+1) = \mathbf{P}_i(\tau_i < \infty) \mathbf{P}_i(\tau_i^{(n)} < \infty) = \alpha \mathbf{P}_i(N > n).$$

Etape 2. On conclut.

- 1 implique 2 : Si i est récurrent, alors $\alpha = 1$ et donc par l'étape 1, on a $N = \infty$ p.s. puisque

$$\mathbf{P}_i(N = \infty) = \lim_{n \rightarrow \infty} \mathbf{P}_i(N > n) = 1,$$

- 2 implique 3 : c'est évident, puisque $G(i, i) = \mathbf{E}_i(N)$.
- 3 implique 1 : Si $G(i, i) = \mathbf{E}_i(N) = \infty$, alors forcément $\alpha = 1$, donc i est récurrent (rappelons que $\alpha = \mathbf{P}_i(\tau_i < \infty)$). En effet, si $\alpha < 1$, alors

$$\mathbf{E}_i(N) = \sum_{n \geq 0} \mathbf{P}_i(N > n) = \sum_{n \geq 0} \alpha^n < \infty,$$

ce qui termine la preuve. □

Remarque. Voici une preuve plus formelle de l'égalité (★): considérons l'application $\Phi : E^{\mathbb{N}} \mapsto \mathbb{N} \cup \{\infty\}$ définie par

$$\Phi((x_k)_{k \geq 0}) = \inf\{k \geq 1 : x_k = i\},$$

et $\Psi : E^{\mathbb{N}} \mapsto \{0, 1\}$ par $\Psi((x_k)_{k \geq 0}) = \mathbf{1}_{\{\Phi((x_k)_{k \geq 0}) < \infty\}}$. On observe alors que

$$\tau_i = \Phi((X_k)_{k \geq 0}) \quad \text{et} \quad \tau_i^{(n+1)} - \tau_i^{(n)} = \Phi((X_{\tau_i^{(n)} + k})_{k \geq 0})$$

donc

$$\mathbf{1}_{\{\tau_i < \infty\}} = \Psi((X_k)_{k \geq 0}) \quad \text{et} \quad \mathbf{1}_{\{\tau_i^{(n+1)} - \tau_i^{(n)} < \infty\}} = \Psi((X_{\tau_i^{(n)} + k})_{k \geq 0}).$$

Ainsi,

$$\begin{aligned} \mathbf{1}_{\{\tau_i^{(n)} < \infty\}} \mathbf{P}_i \left(\tau_i^{(n+1)} - \tau_i^{(n)} < \infty \mid \mathcal{F}_{\tau_i^{(n)}} \right) &= \mathbf{1}_{\{\tau_i^{(n)} < \infty\}} \mathbf{E}_i \left[\Psi((X_{\tau_i^{(n)} + k})_{k \geq 0}) \mid \mathcal{F}_{\tau_i^{(n)}} \right] \\ &= \mathbf{1}_{\{\tau_i^{(n)} < \infty\}} \mathbf{E}_{X_{\tau_i^{(n)}}} \left[\Psi((X_k)_{k \geq 0}) \right] \end{aligned}$$

par Markov forte, soit encore, comme $X_{\tau_i^{(n)}} = i$ (quand $\tau_i^{(n)} < \infty$),

$$\mathbf{1}_{\{\tau_i^{(n)} < \infty\}} \mathbf{P}_i \left(\tau_i^{(n+1)} - \tau_i^{(n)} < \infty \mid \mathcal{F}_{\tau_i^{(n)}} \right) = \mathbf{1}_{\{\tau_i^{(n)} < \infty\}} \mathbf{E}_i [\mathbf{1}_{\{\tau_i < \infty\}}] = \alpha \mathbf{1}_{\{\tau_i^{(n)} < \infty\}}.$$

Théorème 4.10. Soit i un état récurrent (s'il en existe un) de la chaîne de Markov $(X_n)_{n \geq 0}$. Alors, sous $\mathbf{P}_i$, la suite des temps $(\tau_i^{(n)})_{n \geq 0}$ fait de $(X_n)_{n \geq 0}$ un processus régénératif.

Preuve. Il s'agit de montrer que pour tout $f : E \rightarrow \mathbb{R}^+$, les variables aléatoires

$$Z_n = \sum_{k=\tau_i^{(n)}}^{\tau_i^{(n+1)}-1} f(X_k)$$

sont indépendantes et de même loi.

Fixons $n \geq 0$. Comme i est récurrent, on sait que $\tau_i^{(n)} < \infty$ $\mathbf{P}_i$ -p.s. et on a bien sûr $X_{\tau_i^{(n)}} = i$. Par le corollaire 4.8 (et comme $\tau_i^{(n)} < \infty$ p.s.), on sait que $(\tilde{X}_k)_{k \geq 0}$ définie par $\tilde{X}_k = X_{\tau_i^{(n)} + k}$ est une chaîne de même transition P et de loi initiale δ_i , indépendante de $\mathcal{F}_{\tau_i^{(n)}}$. De plus, on voit facilement que $\tilde{\tau}_i = \inf\{k \geq 1 : \tilde{X}_k = i\}$ n'est autre que $\tau_i^{(n+1)} - \tau_i^{(n)}$. Ainsi, on peut écrire

$$Z_n = \sum_{k=0}^{\tau_i^{(n+1)} - \tau_i^{(n)} - 1} f(X_{\tau_i^{(n)} + k}) = \sum_{k=0}^{\tilde{\tau}_i - 1} f(\tilde{X}_k)$$

et il est clair que Z_n a même loi que $Z_1 = \sum_{k=0}^{\tau_i - 1} f(X_k)$.

Enfin, comme les variables aléatoires $Z_1, \dots, Z_{n-1}$ sont toutes $\mathcal{F}_{\tau_i^{(n)}}$ -mesurables, on voit que Z_n est indépendant de $(Z_1, \dots, Z_{n-1})$. $\square$

On déduit donc directement du théorème 3.3:

Corollaire 4.11. Soit i un état récurrent (s'il en existe un) de la chaîne de Markov $(X_n)_{n \geq 0}$. Pour toutes fonctions $f, g : E \rightarrow \mathbb{R}^+$, $\mathbf{P}_i$ -presque sûrement,

$$\lim_{n \rightarrow +\infty} \frac{\sum_{k=0}^n f(X_k)}{\sum_{k=0}^n g(X_k)} = \frac{\int_E f \, d m_i}{\int_E g \, d m_i}$$

si $0 < \int_E g \, d m_i < \infty$, où m_i est la mesure sur E définie par

$$m_i(j) = \mathbf{E}_i \left[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_k=j\}} \right] \quad \forall j \in E.$$

4.5 Mesures et probabilités invariantes

Dans la théorie des chaînes et des processus de Markov, la notion de probabilité invariante est sans doute la plus importante. Elle généralise celle d'état d'équilibre en physique.

Définition. On dit qu'une mesure m sur E est une mesure invariante de la chaîne de Markov de noyau de transition P si $mP = m$, i.e. si pour tout $j \in E$,

$$m(j) = \sum_{i \in E} m(i)P(i, j).$$

Remarque. *i)* Évidemment, la mesure nulle sur E est toujours invariante. Les seules mesures invariantes qui ont un intérêt sont les mesures non triviales.

ii) Si m est une probabilité et si X_0 a pour loi m , alors X_n a pour loi $mP^n = mP(P^{n-1}) = mP^{n-1} = \dots = mP = m$. Donc la loi de la chaîne est *invariante*, c'est la même pour tous les temps.

Les mesures introduites au corollaire 4.11 vont avoir un rôle important.

Lemme 4.12. Soit i un état récurrent de la chaîne de Markov $(X_n)_{n \geq 0}$ (s'il en existe un). Alors, la formule

$$m_i(j) = \mathbf{E}_i \left[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_k=j\}} \right] \quad \forall j \in E$$

définit une mesure invariante.

Preuve. Sous $\mathbf{P}_i$, $X_0 = X_{\tau_i} = i$, on peut donc écrire (distinguer les cas $j = i$ et $j \neq i$)

$$m_i(j) = \mathbf{E}_i \left[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_{k+1}=j\}} \right] = \sum_{k \geq 0} \mathbf{E}_i [\mathbf{1}_{\{\tau_i > k\}} \mathbf{1}_{\{X_{k+1}=j\}}].$$

Comme $\{\tau_i > k\} \in \mathcal{F}_k$, par Markov simple,

$$m_i(j) = \sum_{k \geq 0} \mathbf{E}_i [\mathbf{1}_{\{\tau_i > k\}} \mathbf{P}_i(X_{k+1} = j \mid \mathcal{F}_k)] = \sum_{k \geq 0} \mathbf{E}_i [\mathbf{1}_{\{\tau_i > k\}} \mathbf{P}_{X_k}(X_1 = j)].$$

Mais $P_\ell(X_1 = j) = P(\ell, j)$, donc

$$m_i(j) = \sum_{k \geq 0} \mathbf{E}_i [\mathbf{1}_{\{\tau_i > k\}} P(X_k, j)] = \sum_{\ell \in E} P(\ell, j) \sum_{k \geq 0} \mathbf{E}_i [\mathbf{1}_{\{\tau_i > k\}} \mathbf{1}_{\{X_k=\ell\}}].$$

Comm $\sum_{k \geq 0} \mathbf{E}_i[\mathbf{1}_{\{\tau_i > k\}} \mathbf{1}_{\{X_k = \ell\}}] = \mathbf{E}_i[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_k = \ell\}}] = m_i(\ell)$, on a bien

$$m_i(j) = \sum_{\ell \in E} P(\ell, j) m_i(\ell),$$

donc la mesure m_i est invariante. □

Définition. Une chaîne de Markov de transition P est dite **irréductible** si pour tout $i, j \in E$, il existe $n \in \mathbb{N}$ tel que $P^n(i, j) > 0$.

Autrement dit, pour toute paire d'états $i, j \in E$, il est possible d'aller de i à j

Lemme 4.13. Si m est une mesure invariante d'une chaîne irréductible, alors on a $m(j) > 0$ pour tout $j \in E$ (ou $m(j) = 0$ pour tout $j \in E$).

Preuve. Supposons donc qu'il existe i tel que $m(i) > 0$. Soit ensuite $j \in E$. On sait qu'il existe un entier $n \in \mathbb{N}$ tel que $P^n(i, j) > 0$. Du coup,

$$m(j) = (mP^n)(j) = \sum_{k \in E} m(k) P^n(k, j) \geq m(i) P^n(i, j) > 0.$$

D'autre part, la mesure identiquement nulle est évidemment toujours invariante, quelque soit P . □

Lemme 4.14. Si une chaîne de Markov irréductible possède une **probabilité** invariante π , toute autre mesure invariante est proportionnelle à π . En particulier, π est la seule probabilité invariante.

Remarque. L'hypothèse d'irréductibilité est cruciale : par exemple, supposons que $E = \{1, 2, 3\}$ et que $P(1, 2) = P(2, 1) = P(3, 3) = 1$ (toutes les autres probabilités de transition étant nulles). Cette chaîne n'est pas irréductible (exercice) et on a plusieurs probabilités invariantes, par exemple $m = (1/2 \ 1/2 \ 0)$ (notation matricielle) et $m = (0 \ 0 \ 1)$.

Preuve. Soit m une mesure invariante et i un état fixé de E . Posons $\lambda = \frac{\pi(i)}{m(i)}$ et $m' = \lambda m$ et montrons que $m' = \pi$.

Soit $\mu(k) = \min\{m'(k), \pi(k)\}$. Alors μ est invariante, car $\mu P(k) \leq m' P(k) = m'(k)$ et $\mu P(k) \leq \pi P(k) = \pi(k)$, donc $\mu P(k) \leq \mu(k)$ (pour tout $k \in E$). Comme de plus $\sum_{k \in E} \mu P(k) = \sum_{k \in E} \mu(k)$ (car P est une matrice de transition), on conclut que $\mu P(k) = \mu(k)$ pour tout $k \in E$.

Donc $\pi - \mu$ est une mesure (positive), invariante, et on a $(\pi - \mu)(i) = 0$ (par définition, $m'(i) = \pi(i)$) puis $\mu(i) = \pi(i)$ donc, par le lemme 4.13 $(\pi - \mu)(j) = 0$ pour tout $j \in E$. Autrement dit, $\pi = \mu$.

De même, $m' - \mu$ est une mesure (positive), invariante, et on a $(m' - \mu)(i) = 0$. Donc, par le lemme $(m' - \mu)(j) = 0$ pour tout $j \in E$. Autrement dit, $m' = \mu$.

On a donc bien $m' = \pi$, ce qui conclut la preuve. □

Lemme 4.15 (Principe du maximum). Pour tous $i, j \in E$, $G(i, j) \leq G(j, j)$.

Preuve. On suppose $i \neq j$. Comme $X_k \neq j$ pour tout $k < \tau_j$, on a

$$\begin{aligned} G(i, j) &= \mathbf{E}_i \left[\sum_{k=0}^{\infty} \mathbf{1}_{\{X_k=j\}} \right] \\ &= \mathbf{E}_i \left[\mathbf{1}_{\{\tau_j < \infty\}} \sum_{k=0}^{\infty} \mathbf{1}_{\{X_{k+\tau_j}=j\}} \right] \\ &= \mathbf{E}_i \left[\mathbf{1}_{\{\tau_j < \infty\}} \mathbf{E}_i \left(\sum_{k=0}^{\infty} \mathbf{1}_{\{X_{k+\tau_j}=j\}} \mid \mathcal{F}_{\tau_j} \right) \right]. \end{aligned}$$

Mais par Markov forte, sur l'évènement $\{\tau_j < \infty\}$, on a $\mathbf{E}_i(\sum_{k=0}^{\infty} \mathbf{1}_{\{X_{k+\tau_j}=j\}} \mid \mathcal{F}_{\tau_j}) = \mathbf{E}_{X_{\tau_j}}(\sum_{k=0}^{\infty} \mathbf{1}_{\{X_k=j\}})$, qui n'est autre que $G(j, j)$ puisque $X_{\tau_j} = j$. Ainsi,

$$G(i, j) = \mathbf{P}_i(\tau_j < +\infty) G(j, j) \leq G(j, j). \quad \square$$

Théorème 4.16. Considérons une chaîne de Markov $(X_n)_{n \geq 0}$ irréductible. Les conditions suivantes sont équivalentes.

- (i) Il existe un état $i \in E$ tel que $\mathbf{E}_i(\tau_i) < +\infty$.
- (ii) Pour tout état $i \in E$, $\mathbf{E}_i(\tau_i) < +\infty$.
- (iii) La chaîne de Markov possède une probabilité invariante π .

Sous ces conditions la chaîne est dite **récurrente positive**. Sa probabilité invariante π est unique.

Pour tout $k \in E$,

$$\pi(k) = \frac{1}{\mathbf{E}_k(\tau_k)}, \quad (\star)$$

et pour tout $i, k \in E$

$$\pi(k) = \frac{1}{\mathbf{E}_i(\tau_i)} \mathbf{E}_i \left[\sum_{n=0}^{\tau_i-1} \mathbf{1}_{\{X_n=k\}} \right]. \quad (\star\star)$$

Intuitivement, si π est la probabilité invariante d'une chaîne, $\pi(j)$ représente la proportion du temps passé en j , i.e. $\pi(j) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n \mathbf{1}_{\{X_k=j\}}$ (on verra cela plus loin). Donc $(\star)$ n'est pas déraisonnable : en gros, quand on est en j , on met un temps $\mathbf{E}_j[\tau_j]$ à revenir en j , donc on passe en j "tous les" $\mathbf{E}_j[\tau_j]$, donc la proportion du temps passé en j vaut $1/\mathbf{E}_j[\tau_j]$.

Preuve. Bien sûr, (ii) implique (i). Montrons que (i) implique (iii). Soit $i \in E$ tel que $\mathbf{E}_i(\tau_i) < +\infty$. Alors $\tau_i < \infty$ $\mathbf{P}_i$ -p.s. et i est récurrent. Soit donc m_i la mesure invariante associée à i définie dans le lemme 4.12, i.e. la mesure caractérisée par $m_i(k) = \mathbf{E}_i[\sum_{n=0}^{\tau_i-1} \mathbf{1}_{\{X_n=k\}}]$ pour tout $k \in E$. On a $\sum_{k \in E} m_i(k) = \mathbf{E}_i(\tau_i)$, donc

$$\pi(k) = \frac{1}{\mathbf{E}_i(\tau_i)} \mathbf{E}_i \left[\sum_{n=0}^{\tau_i-1} \mathbf{1}_{\{X_n=k\}} \right]$$

est une probabilité invariante.

Montrons maintenant que (iii) implique (ii). Soit π une probabilité invariante. Pour tout $i \in E$,

$$\mathbf{E}_\pi \left[\sum_{n \geq 0} \mathbf{1}_{\{X_n=i\}} \right] = \sum_{n \geq 0} \mathbf{P}_\pi(X_n = i) = \sum_{n \geq 0} \pi(i) = \infty.$$

Donc $\sum_{j \in E} \pi(j) G(j, i) = \mathbf{E}_\pi[\sum_{n \geq 0} \mathbf{1}_{\{X_n=i\}}] = \infty$ puis, par le lemme 4.15,

$$\infty = \sum_{j \in E} \pi(j) G(j, i) \leq \sum_{j \in E} \pi(j) G(i, i) = G(i, i),$$

ce qui implique que $G(i, i) = \infty$ et donc que i est récurrent.

Par le lemme 4.12, on en déduit que la mesure m_i définie par $m_i(k) = \mathbf{E}_i[\sum_{n=0}^{\tau_i-1} \mathbf{1}_{\{X_n=k\}}]$, $k \in E$, est invariante. Par le lemme 4.14 elle est donc proportionnelle à π , et ainsi, $\mathbf{E}_i(\tau_i) = \sum_{k \in E} m_i(k) < \infty$.

On a montré (★★) dans (i) implique (iii), et (★) n'est autre que (★★) dans le cas où $k = i$. $\square$

Théorème 4.17 (Théorème ergodique des CdM récurrentes positives.). Si $(X_n)_{n \geq 0}$ est une chaîne de Markov récurrente positive de probabilité invariante π , pour toute loi initiale ν , et pour toute fonction $f : E \rightarrow \mathbb{R}^+$, $\mathbf{P}_\nu$ -presque sûrement,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \int_E f d\pi. \quad (\star)$$

(On a $\int_E f d\pi = \sum_{i \in E} f(i)\pi(i)$.) De plus, pour tout $i, j \in E$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k(i, j) = \pi(j).$$

Remarque. En choisissant $f = \mathbf{1}_{\{j\}}$, on voit qu'on a bien, pour n'importe quelle loi initiale, $\pi(j) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n \mathbf{1}_{\{X_k=j\}}$: c'est la proportion du temps passé en j .

Preuve. Comme la chaîne est récurrente positive, tous les états sont récurrent. Par conséquent pour tout $i \in E$, pour tout $f, g : E \rightarrow \mathbb{R}^+$, $\mathbf{P}_i$ -p.s.,

$$\lim_{n \rightarrow +\infty} \frac{\sum_{k=0}^n f(X_k)}{\sum_{k=0}^n g(X_k)} = \frac{\int_E f d m_i}{\int_E g d m_i}$$

où $m_i(j) = \mathbf{E}_i[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_k=j\}}]$. Donc, avec $g = 1$, on trouve

$$\lim_{n \rightarrow +\infty} \frac{\sum_{k=0}^n f(X_k)}{n} = \frac{1}{\mathbf{E}_i[\tau_i]} \mathbf{E}_i \left[\sum_{k=0}^{\tau_i-1} f(X_k) \right] = \int_E f d\pi$$

$\mathbf{P}_i$ -p.s. pour tout $i \in E$ (et donc $\mathbf{P}_\nu$ -p.s. pour toute loi initiale ν).

Avec $f = \mathbf{1}_{\{j\}}$, on obtient $\frac{1}{n} \sum_{k=0}^n \mathbf{1}_{\{X_k=j\}} \rightarrow \pi(j)$ $\mathbf{P}_i$ -p.s. Donc par convergence dominée,

$$\frac{1}{n} \sum_{k=0}^n P^k(i, j) = \mathbf{E}_i \left[\frac{1}{n} \sum_{k=0}^n \mathbf{1}_{\{X_k=j\}} \right] \xrightarrow{n \rightarrow \infty} \pi(j). \quad \square$$

Ce théorème donne deux moyens pratiques d'approcher π si, comme c'est souvent le cas, on ne sait pas la calculer explicitement. La première façon est la méthode de Monte Carlo, qui consiste à simuler sur ordinateur une longue trajectoire X_n de la chaîne, et d'utiliser le fait que $\pi(j) \simeq \frac{1}{n} \sum_{k=0}^n \mathbf{1}_{\{X_k=j\}}$. L'autre façon est de calculer itérativement P^n , par exemple dans le cas où E est fini. Puis de faire la moyenne des $P^n(i, j)$ pour approcher $\pi(j)$.

Définition. On dit qu'une suite $(x_n)_{n \geq 0}$ à valeurs dans E (dénumbrable) **tend vers l'infini** si pour toute partie finie F de E , il existe $n_F \geq 1$ t.q. pour tout $n \geq n_F$, $x_n \notin F$.

Théorème 4.18. Pour une chaîne de Markov irréductible, on a l'alternative :

Cas récurrent: Tous les états sont récurrents et partant de tout point, la chaîne visite une infinité de fois tous les autres, p.s.

Cas transitoire: Partant de tout point, $(X_n)_{n \geq 0}$ tend vers l'infini, p.s.

Preuve. S'il existe un état récurrent i , considérons la mesure invariante $m_i(j) = \mathbf{E}_i[\sum_{k=0}^{\tau_i-1} \mathbf{1}_{\{X_k=j\}}]$ qui lui est associée. Soit $j \in E$ fixé. La chaîne étant irréductible, $m_i(j) > 0$. Or, par le théorème ergodique avec $f = \mathbf{1}_{\{j\}}$ et $g = \mathbf{1}_{\{i\}}$ on a $\mathbf{P}_i$ -p.s.,

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=0}^{n-1} \mathbf{1}_{\{X_k=j\}}}{\sum_{k=0}^{n-1} \mathbf{1}_{\{X_k=i\}}} = \frac{m_i(j)}{m_i(i)} > 0.$$

Puisque i est récurrent, $\sum_{k=0}^{\infty} \mathbf{1}_{\{X_k=i\}} = +\infty$. La limite précédente assure donc que $\sum_{k=0}^{\infty} \mathbf{1}_{\{X_k=j\}} = +\infty$ $\mathbf{P}_i$ -p.s. En prenant l'espérance, on voit que $G(i, j) = \infty$, d'où $G(j, j) = +\infty$ par le principe du maximum : l'état j est donc récurrent.

Si au contraire il n'y a pas d'état récurrent, alors pour tout $j \in E$, $G(j, j) < \infty$. Par conséquent si F est une partie finie de E ,

$$\mathbf{E}_i \left[\sum_{k=0}^{\infty} \mathbf{1}_F(X_k) \right] = \sum_{j \in F} G(i, j) \leq \sum_{j \in F} G(j, j) < +\infty.$$

Donc $\sum_{k=0}^{\infty} \mathbf{1}_F(X_k)$ est une somme finie, $\mathbf{P}_i$ -p.s., ce qui n'est possible que si X_n quitte F définitivement après un certain temps. Comme F est arbitraire, $(X_n)_{n \geq 0}$ tends vers l'infini $\mathbf{P}_i$ -p.s. $\square$

Corollaire 4.19. Si E est fini, toute chaîne irréductible est récurrente (car ne peut pas tendre vers l'infini) positive (car admet donc une mesure invariante qui est forcément finie).

Exemple 4.3. Avec $E = \{1, 2, 3\}$ et $P(2, 1) = P(2, 3) = 1/2$ et $P(1, 2) = P(3, 2) = 1$. Faire un dessin. Comprendre qu'intuitivement, $\pi = (1/4 \ 1/2 \ 1/4)$. Calculer π en résolvant $\pi = \pi P$. Calculer $\mathbf{E}_i[\tau_i]$ (à la main) pour $i = 1, 2, 3$ et vérifier qu'on a bien $\pi(i) = 1/\mathbf{E}_i(\tau_i)$.

Proposition 4.20. Une chaîne irréductible **récurrente** admet une et une seule (à constante multiplicative près) mesure invariante.

Preuve. L'existence d'une mesure invariante m a été vue au lemme 4.12. Pour l'unicité, on procède comme au lemme 4.14 : on considère deux mesures invariantes m et ℓ , toutes deux non identiquement nulles, donc strictement positives partout par le lemme 4.13. On fixe $i \in E$ et on pose $m' = \lambda \ell$, où $\lambda = m(i)/\ell(i)$.

On vérifie que $\mu(j) = \min\{m(j), m'(j)\}$ est invariante. En effet, on a bien sûr $\mu P(j) \leq \mu(j)$ (comme au lemme 4.14). On considère ensuite la mesure positive $\nu(j) = \mu(j) - \mu P(j)$, et on écrit, pour $n \geq 1$,

$$\sum_{k=0}^n \nu P^k = \sum_{k=0}^n \mu P^k - \mu P^{k+1} = \mu - \mu P^{n+1} \leq \mu.$$

Donc $\nu G = \sum_{k=0}^{\infty} \nu P^k \leq \mu$, i.e. pour tout $j \in E$, $\sum_{i \in E} \nu(i) G(i, j) \leq \mu(j)$. C'est impossible si ν n'est pas identiquement nulle, puisque $G(i, j) = \infty$ pour tout $i, j \in E$ (rappelons que la chaîne est supposée récurrente). Nous avons montré que $\mu P = \mu$, qui est donc invariante.

Donc $m - \mu$ est une mesure (positive), invariante, nulle en i (par construction), donc identiquement nulle (cf lemme 4.13). Donc $m = \mu$. On montre de même que $m' = \mu$, ce qui implique donc que $m = \mu = m'$, c'est-à-dire que les mesures m et ℓ sont proportionnelles. $\square$

Exercice 4.3. Considérons la marche aléatoire simple sur $\mathbb{Z}$, i.e. $P(i, i+1) = p$, $P(i, i-1) = 1-p$, avec $p \in]0, 1[$. Montrer qu'elle est irréductible. Montrer que $m(i) = 1$ pour tout i est invariante (quelque soit la valeur de p).

A l'aide de la loi des grands nombres (écrire X_n comme une somme de v.a. i.i.d.), montrer que $(X_n)_{n \geq 0}$ est transitoire si $p \neq 1/2$.

Si e.g. $p > 1/2$, montrer que les mesures $m_i = 1$ pour tout $i \in \mathbb{Z}$ et $m_i = [p/(1-p)]^i$ pour tout $i \in \mathbb{Z}$ sont invariantes.

Si $p = 1/2$, montrer que $\mathbf{P}_0(X_{2n} = 0) = C_{2n}^n 2^{-2n} \sim 1/\sqrt{\pi n}$. Montrer que $G(0, 0) = \infty$. Donc $(X_n)_{n \geq 0}$ est récurrente. Mais non récurrente positive (sinon, elle aurait une probabilité invariante, $m \equiv 1$ lui serait proportionnelle, ce qui est absurde). Remarquer aussi que, comme $m(i) = \mathbf{E}_0[\sum_{k=0}^{\tau_0-1} \mathbf{1}_{\{X_k=i\}}]$ est invariante (et vérifie $m(0) = 1$), on a donc $m(i) = 1$ pour tout $i \in \mathbb{Z}$. C'est très surprenant : le nombre moyen de passages par $i \in \mathbb{Z}$ entre deux passages en 0 vaut toujours 1, que $i = 1$ ou que $i = 10000$!!

Exemple 4.4 (Une chaîne de Markov sans mesure invariante). On considère la chaîne sur $\mathbb{N}$ avec les transitions $P(i, i+1) = p_i$ et $P(i, 0) = 1 - p_i$ pour $i \geq 0$. On choisit la suite $(p_i)_{i \geq 0}$ de façon que $p_0 = 1$, $p_i > 0$ pour tout $i \geq 1$ et $\sum_{i \geq 1} (1 - p_i) < 1$. Alors, la chaîne est irréductible. De plus, une mesure invariante π doit satisfaire

$$\pi_0 = \sum_{i \geq 1} \pi_i (1 - p_i) \quad \text{et} \quad \pi_{i+1} = p_i \pi_i \quad \text{pour } i \geq 0 \quad (\star)$$

En particulier, $\pi_i = \pi_0 \prod_{j=1}^{i-1} p_j$, ce qui implique, si $\pi_0 > 0$,

$$\sum_{i \geq 1} (1 - p_i) \pi_i = \pi_0 \sum_{i \geq 1} (1 - p_i) \prod_{j=1}^{i-1} p_j \leq \pi_0 \sum_{i \geq 1} (1 - p_i) < \pi_0,$$

par hypothèse. Par conséquent, la première condition de $(\star)$ ne peut être satisfaite que si $\pi_0 = 0$, mais alors $\pi_i = 0$ pour tout $i \geq 0$, et la mesure est triviale. D'après le Lemme 4.12, puisque $m_i(i) = 1$, la chaîne doit donc être transiente (le vérifier directement!).

Résumé important. Il est important de garder en tête que

- (i) Si une chaîne irréductible a une probabilité invariante, alors elle est récurrente positive (et la probabilité invariante est unique) cf Théorème 4.16.
- (ii) Il est possible qu'une chaîne irréductible admette une (ou plusieurs) mesures invariantes sans qu'elle soit récurrente cf Exemple 4.3.
- (iii) Si une chaîne irréductible est récurrente, alors elle admet une unique (à constante près) mesure invariante, cf Proposition 4.20.
- (iv) Une chaîne irréductible peut ne pas admettre de mesure invariante non-triviale, elle alors être transiente, cf Exemple 4.4.

Théorème 4.21. Soit $(X_n)_{n \geq 0}$ une marche aléatoire à valeurs dans $\mathbb{Z}$, de pas de loi μ (une probabilité sur $\mathbb{Z}$). C'est bien sûr une chaîne de transition $P(i, j) = \mu(j - i)$. Si μ est centrée, i.e. si $\sum_{k \in \mathbb{Z}} k\mu(k) = 0$, alors tout état est récurrent.

La preuve est donnée dans l'appendice (théorème B.1). Bien sûr, cette chaîne est transitoire dans le cas où $\sum_{k \in \mathbb{Z}} k\mu(k) \neq 0$, il suffit d'utiliser la loi des grands nombres.

4.6 Réversibilité

Proposition 4.22. Soit $(X_n)_{n \geq 0}$ une chaîne de Markov récurrente positive de probabilité invariante ν . Sous $\mathbf{P}_\nu$, le processus $(X_{N-n})_{n \in \{0, \dots, N\}}$ (avec N grand fixé) est une chaîne de Markov de transition.

$$Q(i, j) = \frac{\nu(j)}{\nu(i)} P(j, i)$$

Preuve. Pour tout $n \in \{0, \dots, N\}$ et $a_0, a_1, \dots, a_n \in E$, on a:

$$\begin{aligned} \mathbf{P}_\nu(X_N = a_0, X_{N-1} = a_1, \dots, X_{N-n} = a_n) \\ &= \mathbf{P}_\nu(X_{N-n} = a_n, X_{N-n+1} = a_{n-1}, \dots, X_N = a_0) \\ &= \nu P^{N-n}(a_n) P(a_n, a_{n-1}) \cdots P(a_1, a_0) \\ &= \nu(a_n) P(a_n, a_{n-1}) \cdots P(a_1, a_0) \\ &= \frac{\nu(a_n)}{\nu(a_{n-1})} P(a_n, a_{n-1}) \frac{\nu(a_{n-1})}{\nu(a_{n-2})} P(a_{n-1}, a_{n-2}) \cdots \frac{\nu(a_1)}{\nu(a_0)} P(a_1, a_0) \nu(a_0) \\ &= Q(a_{n-1}, a_n) Q(a_{n-2}, a_{n-1}) \cdots Q(a_0, a_1) \nu(a_0) \end{aligned}$$

ce qui montre la proposition. On a utilisé que la loi de X_{N-n} est νP^{N-n} et que $\nu P^{N-n} = \nu$ car ν est invariante. $\square$

Définition. Une probabilité ν sur E est dite réversible (ou P -réversible) si pour tout $i, j \in E$

$$\nu(i) P(i, j) = \nu(j) P(j, i).$$

Remarque. (i) Une probabilité réversible ν est invariante, pour s'en rendre compte, il suffit de sommer sur j :

$$\sum_{j \in E} \nu(j) P(j, i) = \sum_{j \in E} \nu(i) P(i, j) = \nu(i).$$

D'ailleurs, une des méthodes les plus facile pour montrer qu'une mesure est invariante et de montrer qu'elle est en fait réversible, voir l'exemple suivant.

(ii) L'équation de la définition est parfois appelée "équilibre local" (on a une identité par paire d'états, qui implique la stationarité, qui est globale). Évidemment, la réversité n'est possible que si $P(i, j) > 0$ ssi $P(j, i) > 0$.

Exemple 4.5 (Marche aléatoire sur un graphe). On considère un graphe G sur un ensemble fini de sommets V , c'est-à-dire un couple $G = (V, E)$ où E est un sous-ensemble de l'ensemble des paires de sommets. Pour simplifier les notations, on note $u \sim v$ si $\{u, v\} \in E$. Le degré d'un sommet $u \in V$ est le nombre d'arêtes qui lui sont adjacentes, c'est-à-dire $\deg(u) = \#\{v \in V : v \sim u\}$. On suppose que le graphe G est connexe, c'est-à-dire qu'il existe un chemin d'arêtes reliant n'importe quelle paire de sommets u et v .

La marche aléatoire simple sur le graphe G est la chaîne de Markov de matrice de transition P sur V donnée par $P(u, v) = \mathbf{1}_{\{v \sim u\}} / \deg(u)$, $u, v \in V$: si on est à un sommet u , à l'étape suivante on choisit uniformément un voisin de u . Comme le graphe est connexe, la chaîne est irréductible.

La mesure ν sur V donnée par $\nu(u) = \deg(u)$ est une mesure réversible pour P , et donc une mesure invariante. En effet pour $u, v \in V$,

$$\nu(u)P(u, v) = \deg(u) \cdot \frac{1}{\deg(u)} \mathbf{1}_{\{u \sim v\}} = \mathbf{1}_{\{u \sim v\}} = \deg(v) \cdot \frac{1}{\deg(v)} \mathbf{1}_{\{v \sim u\}} = \nu(v)P(v, u).$$

En particulier, si le graphe est *régulier* c'est-à-dire si les degrés de tous les sommets sont identiques, alors la mesure constante est invariante.

4.7 Processus de naissance et mort à temps discret

Dans cette section, on étudie un cas particulier de chaînes, mais les méthodes utilisées (pour montrer la transience ou la récurrence) sont assez générales.

Définition. Une chaîne de Markov à temps discret à valeurs dans $E = \mathbb{N}$ est appelée processus de naissance et mort (PNM) à temps discret si $P(i, j) = 0$ dès que $|i - j| \neq 1$. On pose alors $p_i = P(i, i + 1)$ (pour $i \geq 0$) et $q_i = P(i, i - 1)$ (pour $i \geq 1$). On a forcément $p_0 = 1$ et $p_i + q_i = 1$ pour $i \geq 1$.

Remarque. Un processus de naissance et mort à temps discret est irréductible si et seulement si $p_i > 0$ et $q_i > 0$ pour tout $i \geq 1$.

Théorème 4.23. Un PNM à temps discret irréductible est

(i) récurrent si et seulement si

$$\sum_{i \geq 1} \frac{q_1 q_2 \dots q_i}{p_1 p_2 \dots p_i} = \infty;$$

(ii) récurrent positif si et seulement si

$$S = 1 + \frac{1}{q_1} + \sum_{i \geq 2} \frac{p_1 p_2 \dots p_{i-1}}{q_1 q_2 \dots q_i} < \infty.$$

Dans ce dernier cas, son unique probabilité invariante est alors donnée par $\pi_0 = 1/S$, $\pi_1 = 1/(Sq_1)$ et, pour tout $i \geq 2$, $\pi_i = (p_1 p_2 \dots p_{i-1}) / (Sq_1 q_2 \dots q_i)$.

Preuve. Commençons par (ii). Une mesure m est invariante si et seulement si $mP = m$, ce qui se réécrit

$m_0 = m_1 q_1$ et $m_i = m_{i+1} q_{i+1} + m_{i-1} p_{i-1}$ pour $i \geq 1$, i.e.

$$m_{i+1} = \frac{m_i - m_{i-1} p_{i-1}}{q_{i+1}}.$$

Cette récurrence d'ordre 2 se résout ainsi :

$$m_1 = m_0 / q_1, \quad \text{et pour } i \geq 2, \quad m_i = m_0 \frac{p_1 p_2 \dots p_{i-1}}{q_1 q_2 \dots q_i}.$$

Le point (ii) s'ensuit aisément (rappelons qu'une chaîne est récurrente positive si et seulement si elle admet une probabilité invariante).

Point (i), étape 1. La chaîne est récurrente si et seulement si $\mathbf{P}_1(T_0 < \infty) = 1$, où $T_0 = \inf\{n \geq 0 : X_n = 0\}$. On observe d'abord que, la chaîne étant irréductible, il suffit de considérer la récurrence d'un état quelconque; nous prendrons l'état 0. Maintenant, on a $\mathbf{P}_0(\tau_0 < \infty) = \mathbf{P}_1(T_0 < \infty)$ par Markov simple: en notant que $X_1 = 1$ sous $\mathbf{P}_0$, en introduisant $\tilde{X}_n = X_{n+1}$ et en notant que $\tau_0 = 1 + \tilde{T}_0$ (où $\tilde{T}_0 = \inf\{n \geq 0 : \tilde{X}_n = 0\}$),

$$\begin{aligned} \mathbf{P}_0(\tau_0 < \infty) &= \mathbf{E}_0[\mathbf{1}_{\{X_1=1\}} \mathbf{P}_0(\tau_0 < \infty \mid \mathcal{F}_1)] \\ &= \mathbf{E}_0[\mathbf{1}_{\{X_1=1\}} \mathbf{P}_0(\tilde{T}_0 < \infty \mid \mathcal{F}_1)] \\ &= \mathbf{E}_0[\mathbf{1}_{\{X_1=1\}} \mathbf{P}_1(T_0 < \infty)] \\ &= \mathbf{P}_0(X_1 = 1) \mathbf{P}_1(T_0 < \infty) \\ &= \mathbf{P}_1(T_0 < \infty). \end{aligned}$$

Point (i), étape 2. On fixe $a \in \mathbb{N}_*$ (grand). On pose $T_a = \inf\{n \geq 0 : X_n = a\}$ et $u(i) = \mathbf{P}_i(T_0 < T_a)$ pour $i = 0, \dots, a$. On a bien sûr $u(0) = 1$ et $u(a) = 0$, et

$$u(i) = q_i u(i-1) + p_i u(i+1) \quad \text{pour } i = 1, \dots, a-1.$$

En effet, en posant $\tilde{X}_n = X_{n+1}$ et $\tilde{T}_j = \inf\{n \geq 0 : \tilde{X}_n = j\}$, on a, pour $i = 1, \dots, a-1$, $T_0 = \tilde{T}_0 + 1$ et $T_a = \tilde{T}_a + 1$ et donc par Markov simple

$$\begin{aligned} \mathbf{P}_i[T_0 < T_a] &= \mathbf{E}_i[\mathbf{P}_i(T_0 < T_a \mid \mathcal{F}_1)] \\ &= \mathbf{E}_i[\mathbf{P}_i(\tilde{T}_0 < \tilde{T}_a \mid \mathcal{F}_1)] \\ &= \mathbf{E}_i[\mathbf{P}_{X_1}(T_0 < T_a)] \\ &= \mathbf{P}_i(X_1 = i-1) \mathbf{P}_{i-1}(T_0 < T_a) + \mathbf{P}_i(X_1 = i+1) \mathbf{P}_{i+1}(T_0 < T_a) \\ &= q_i \mathbf{P}_{i-1}(T_0 < T_a) + p_i \mathbf{P}_{i+1}(T_0 < T_a). \end{aligned}$$

Point (i), étape 3. On a donc, pour $i = 1, \dots, a-1$, $p_i(u(i+1) - u(i)) = q_i(u(i) - u(i-1))$, et donc $u(i+1) - u(i) = (q_i / p_i) \dots (q_1 / p_1) [u(1) - u(0)]$, puis

$$u(i) = u(0) + \sum_{k=0}^{i-1} (u(k+1) - u(k)) = u(0) + [u(1) - u(0)] \cdot \left(1 + \sum_{k=1}^{i-1} \frac{q_1 \dots q_k}{p_1 \dots p_k} \right).$$

En choisissant $i = a$, comme $u(a) = 0$ et $u(0) = 1$, on trouve

$$\mathbf{P}_1(T_0 < T_a) = u(1) = 1 - \frac{1}{1 + \sum_{k=1}^{a-1} \frac{q_1 \dots q_k}{p_1 \dots p_k}}.$$

Point (i), étape 4. Mais sous $\mathbf{P}_1$, on a

$$\{T_0 < \infty\} = \bigcup_{a \geq 1} \{T_0 < T_a\} \quad (\text{union croissante}),$$

car $T_a < \infty$ $\mathbf{P}_1$ -p.s. pour tout $a \in \mathbb{N}_*$ (car c'est évident si la chaîne est récurrente, et c'est vrai aussi si la chaîne est transitoire, puisqu'alors $X_n \rightarrow \infty$ et est donc forcée de passer par a en partant de 1) ; et car $T_a \geq a - 1 \rightarrow \infty$ $\mathbf{P}_1$ -p.s. quand $a \rightarrow \infty$.

Finalement, en mettant les résultats des étapes ensemble

$$\mathbf{P}_0(\tau_0 < \infty) = \mathbf{P}_1(T_0 < \infty) = \lim_{a \rightarrow \infty} \mathbf{P}_1(T_0 < T_a)$$

et $\mathbf{P}_0(\tau_0 < \infty) = 1$ si et seulement si $\sum_{k=1}^{\infty} \frac{q_1 \dots q_k}{p_1 \dots p_k} = \infty$. □

4.8 Problèmes d'absorption

Considérons une chaîne de Markov $(X_n)_{n \geq 0}$ d'espace d'état fini E , de transition P . On suppose que $E = F \cup G$ (avec F et G disjoints) et que la transition a les propriétés suivantes:

- (a) pour tout $i \in G$, $P(i, i) = 1$,
- (b) pour tout $i \in F$, il existe $j \in G$ et $n \geq 1$ t.q. $P^n(i, j) > 0$.

Noter que la matrice s'écrit donc sous la forme

$$P = \begin{pmatrix} Q & A \\ 0 & I \end{pmatrix}.$$

Ces matrices sont toutes de natures différentes : $(Q(i, j))_{i \in F, j \in F}$, $(A(i, j))_{i \in F, j \in G}$ et $(I(i, j))_{i \in G, j \in G}$.

Noter qu'une telle chaîne n'est pas du tout irréductible. Les éléments de G sont appelés états absorbants (une fois qu'on y est, on y reste). L'énoncé suivant permet de calculer, partant de $i \in F$, la probabilité de finir absorbé en $j \in G$, ainsi que d'autres quantités tout à fait intéressantes.

Proposition 4.24. Sous les conditions (a) et (b) plus haut:

- (i) La matrice $N = \sum_{k=0}^{+\infty} Q^k$ existe et vérifie $N = (I - Q)^{-1}$.
- (ii) On a $\lim_{n \rightarrow +\infty} P^n = \begin{pmatrix} 0 & NA \\ 0 & I \end{pmatrix}$.
- (iii) Soit $T_G = \inf\{n \geq 0 : X_n \in G\}$. Pour tout $i \in F$ et $j \in G$, on a

$$\mathbf{P}_i(X_{T_G} = j) = (NA)(i, j) \quad \text{et} \quad \mathbf{E}_i[T_G] = \sum_{j \in G} N(i, j).$$

Preuve. Remarquons d'abord (récurrence) que

$$P^n = \begin{pmatrix} Q^n & (I + Q + Q^2 + \dots + Q^{n-1})A \\ 0 & I \end{pmatrix}.$$

Pour tout $i \in F$, i est transitoire. En effet, il suffit de montrer que $\mathbf{P}_i(\tau_i < \infty) < 1$. Mais, par (b), on sait qu'il existe $j \in G$ et $n \geq 1$ t.q. $P^n(i, j) > 0$, ce qui implique qu'on peut trouver $i_1, \dots, i_{n-1} \in F$ tels que

$P(i, i_1)P(i_1, i_2) \cdots P(i_{n-1}, j) > 0$. Quitte à raccourcir ce chemin, on peut supposer que tous les i_k sont différents de i . Du coup,

$$\begin{aligned} \mathbf{P}_i(\tau_i = \infty) &\geq \mathbf{P}_i(X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = j) \\ &= P(i, i_1)P(i_1, i_2) \cdots P(i_{n-1}, j) > 0. \end{aligned}$$

Par conséquent, pour tout $i \in F$, $G(i, i) < +\infty$. Il résulte donc du principe du maximum que $G(i, j) < +\infty$ pour tout $i, j \in F$. Autrement dit la matrice N est finie (observer que $P^n(i, j) = Q^n(i, j)$ pour tout $i, j \in F$). En particulier Q^n tend vers 0. Comme $(I - Q)(I + Q + \cdots + Q^n) = I - Q^{n+1}$ on en déduit que $(I - Q)N = I$, soit encore $N = (I - Q)^{-1}$. On conclut aussi que

$$\lim_{n \rightarrow +\infty} P^n = \begin{pmatrix} 0 & NA \\ 0 & I \end{pmatrix}.$$

Notons que $T_G = \sum_{n \geq 0} \sum_{j \in F} \mathbf{1}_{\{X_n = j\}}$. Ainsi, pour tout $i \in F$,

$$\mathbf{E}_i[T_G] = \sum_{n \geq 0} \sum_{j \in F} P^n(i, j) = \sum_{n \geq 0} \sum_{j \in F} Q^n(i, j) = \sum_{j \in F} N(i, j),$$

cette somme (finie) étant bien sûr finie (puisque E est fini). Enfin, une fois dans G la chaîne ne bouge plus, et donc $X_{T_G} = \lim_{n \rightarrow +\infty} X_n$. Ainsi, pour $i \in F$ et $j \in G$,

$$\mathbf{P}_i(X_{T_G} = j) = \lim_{n \rightarrow +\infty} \mathbf{P}_i(X_n = j) = \lim_{n \rightarrow +\infty} P^n(i, j) = (NA)(i, j). \quad \square$$

Exemple 4.6. J'ai 1 euro et j'ai besoin de 3 euros. Je joue à pile ou face avec un ami. Je mise toujours 1, et à chaque étape, j'en récupère donc 0 avec probabilité 1/2 et 2 avec probabilité 1/2. J'arrête quand j'ai 0 ou 3 euros.

Ma fortune X_n à l'instant n est une chaîne de Markov, de loi initiale δ_1 , sur $E = \{0, 1, 2, 3\}$. On a $P(0, 0) = 1$, $P(1, 0) = P(1, 2) = 1/2$, $P(2, 1) = P(2, 3) = 1/2$ et $P(3, 3) = 1$. On peut écrire $E = F \cup G$ avec $F = \{1, 2\}$ et $G = \{0, 3\}$. En réordonnant les états de manière à exhiber les blocs Q , A et I selon F et G , la matrice de transition est alors donnée par

$$\begin{matrix} & \begin{matrix} 1 & 2 & 0 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 0 \\ 3 \end{matrix} & \left(\begin{array}{cc|cc} \hline 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline \end{array} \right) \end{matrix}$$

On peut alors facilement calculer

$$N = (I - Q)^{-1} = \frac{1}{2} \begin{pmatrix} 4/3 & 2/2 \\ 2/3 & 2/3 \end{pmatrix} \quad \text{et} \quad NA = \frac{1}{2} \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}$$

(Attention aux natures de ces matrices). On trouve que $NA(1, 3) = 1/3$: la probabilité, partant d'une fortune de 1 euro, de terminer avec 3 euros, vaut, comme il se doit, $1/3$. L'espérance du temps de jeu est $\mathbf{E}_1(T_G) = N(1, 1) + N(1, 2) = 4/3 + 2/3 = 2$.

Processus markoviens de sauts

Dans toute la suite, E est un ensemble dénombrable, éventuellement fini, qu'on munit de la distance discrète : $d(i, j) = \mathbf{1}_{\{i \neq j\}}$. Nous allons étudier les "processus de Markov à temps continu" à valeurs dans E . On les appelle aussi "chaînes de Markov à temps continu" ou encore "processus markoviens de saut".

5.1 Premières propriétés

Définition (Processus markovien de sauts). Soit $(P_t)_{t \in \mathbb{R}^+}$ une famille de matrices de transition sur E , c'est-à-dire telles que $P_t(i, j) \geq 0$ et $\sum_{j \in E} P_t(i, j) = 1$.

Soit $(X_t)_{t \geq 0}$ à valeurs dans E . On note $\mathcal{F}_s = \sigma(X_u, u \leq s)$. Le processus $(X_t)_{t \geq 0}$ est appelé **processus markovien de saut** (PMS) de semigroupe $(P_t)_{t \geq 0}$ si, pour tout $j \in E$, tout $0 \leq s \leq t + s$,

$$\mathbf{P}(X_{s+t} = j \mid \mathcal{F}_s) = P_t(X_s, j),$$

et si $t \mapsto X_t$ est p.s. continu à droite sur $\mathbb{R}^+$.

Remarque. On rappelle qu'on considère E avec la distance discrète. Une fonction $x : \mathbb{R}_+ \rightarrow E$ est continue à droite pour la distance discrète ssi pour tout $t \geq 0$, il existe $\eta > 0$ tel que pour tout $s \in [t, t + \eta]$, on ait $x(s) = x(t)$.

On peut montrer (à l'aide du théorème de classe monotone) le résultat suivant (voir aussi les compléments au chapitre A)).

Remarque. Pour montrer qu'un processus p.s. continu à droite $(X_t)_{t \geq 0}$ à valeurs dans E est un PMS de semigroupe $(P_t)_{t \geq 0}$, il suffit de montrer que pour tout $n \in \mathbb{N}$, pour tous $0 \leq t_1 \leq \dots \leq t_{n+1}$, pour tout $i_1, \dots, i_{n+1} \in E$,

$$\mathbf{P}(X_{t_{n+1}} = i_{n+1} \mid X_{t_1} = i_1, \dots, X_{t_n} = i_n) = P_{t_{n+1} - t_n}(i_n, i_{n+1}).$$

Donnons un premier exemple de tel processus, qui s'avérera relativement général.

Proposition 5.1. Soit $(Z_n)_{n \in \mathbb{N}}$ une chaîne de Markov sur E de probabilité de transition Q indépendante d'un processus de Poisson $(N_t)_{t \in \mathbb{R}^+}$ de paramètre λ . Alors

$$(X_t = Z_{N_t}, t \in \mathbb{R}^+)$$

est un PMS de semigroupe

$$P_t(i, j) = \sum_{n=0}^{+\infty} e^{-\lambda t} \frac{(\lambda t)^n Q^n(i, j)}{n!}.$$

Preuve. Notons $0 < T_1 < T_2 \dots$ les instants de saut du processus de Poisson. On a alors $X_t = Z_0$ pour $t \in [0, T_1[$, puis $X_t = Z_1$ pour $t \in [T_1, T_2[$, etc. $(X_t)_{t \geq 0}$ est donc bien continu à droite p.s. On écrit ensuite (notons ν la loi de Z_0)

$$\begin{aligned} & \mathbf{P}(X_0 = i_0, X_{t_1} = i_1, \dots, X_{t_r} = i_r) \\ &= \sum_{k_1, \dots, k_r \geq 0} \mathbf{P}\left(N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2, \dots, N_{t_r} - N_{t_{r-1}} = k_r, \right. \\ & \quad \left. Z_0 = i_0, Z_{k_1} = i_1, Z_{k_1+k_2} = i_2, \dots, Z_{k_1+\dots+k_r} = i_r\right) \\ &= \sum_{k_1, \dots, k_r \geq 0} \mathbf{P}\left(N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2, \dots, N_{t_r} - N_{t_{r-1}} = k_r\right) \\ & \quad \times \mathbf{P}\left(Z_0 = i_0, Z_{k_1} = i_1, Z_{k_1+k_2} = i_2, \dots, Z_{k_1+\dots+k_r} = i_r\right) \\ &= \sum_{k_1, \dots, k_r \geq 0} \frac{(\lambda t_1)^{k_1}}{k_1!} e^{-\lambda t_1} \times \dots \times \frac{(\lambda(t_r - t_{r-1}))^{k_r}}{k_r!} e^{-\lambda(t_r - t_{r-1})} \times \nu(i_0) Q^{k_1}(i_0, i_1) \dots Q^{k_r}(i_{r-1}, i_r) \\ &= \nu(i_0) P_{t_1}(i_0, i_1) P_{t_2-t_1}(i_1, i_2) \times \dots \times P_{t_r-t_{r-1}}(i_{r-1}, i_r). \end{aligned}$$

On conclut bien sûr que

$$\begin{aligned} \mathbf{P}(X_{t_r} = i_r \mid X_0 = i_0, X_{t_1} = i_1, \dots, X_{t_{r-1}} = i_{r-1}) &= \frac{\nu(i_0) P_{t_1}(i_0, i_1) P_{t_2-t_1}(i_1, i_2) \times \dots \times P_{t_r-t_{r-1}}(i_{r-1}, i_r)}{\nu(i_0) P_{t_1}(i_0, i_1) P_{t_2-t_1}(i_1, i_2) \times \dots \times P_{t_{r-1}-t_{r-2}}(i_{r-2}, i_{r-1})} \\ &= P_{t_r-t_{r-1}}(i_{r-1}, i_r). \end{aligned} \quad \square$$

Proposition 5.2 (Equation de Chapman–Kolmogorov). Supposons que E soit le “vrai” espace d'états du PMS, c'est à dire que pour tout $i \in E$, il existe $t_i \geq 0$ (déterministe) tel que $\mathbf{P}(X_{t_i} = i) > 0$. Alors $(P_t)_{t \in \mathbb{R}^+}$, d'un PMS a la propriété de semigroupe suivante: pour tous $t, s \geq 0$, $P_{t+s} = P_t P_s$.

Preuve. On écrit

$$\begin{aligned} P_{t+s}(i, j) &= \mathbf{P}(X_{t_i+t+s} = j \mid X_{t_i} = i) \\ &= \sum_{k \in E} \mathbf{P}(X_{t_i+t+s} = j \mid X_{t_i+t} = k, X_{t_i} = i) \mathbf{P}(X_{t_i+t} = k \mid X_{t_i} = i) \\ &= \sum_{k \in E} P_s(k, j) P_t(i, k) = (P_t P_s)(i, j). \end{aligned} \quad \square$$

Notation. Comme dans le cas des chaînes, on note $\mathbf{P}_i, \mathbf{E}_i$ (resp. $\mathbf{P}_v, \mathbf{E}_v$) la loi du processus lorsque $X_0 = i$ (resp. $X_0 \sim v$).

Rappelons qu'un temps d'arrêt T du processus $(X_t)_{t \geq 0}$ est une v.a. à valeurs dans $\mathbb{R}^+ \cup \{\infty\}$ telle que pour tout $t \geq 0$, $\{T \leq t\} \in \mathcal{F}_t$. On introduit alors la tribu $\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T \leq t\} \in \mathcal{F}_t \forall t \geq 0\}$. Comme d'habitude, on a posé préalablement $\mathcal{F}_t = \sigma(\{X_s, s \leq t\})$.

Théorème 5.3 (Propriété de Markov forte). Soit $(X_t)_{t \geq 0}$ un PMS à valeurs dans E et T un temps d'arrêt. Pour toute fonction mesurable $f : E^{\mathbb{R}^+} \rightarrow \mathbb{R}$, bornée ou positive,

$$\mathbf{1}_{\{T < \infty\}} \mathbf{E}[f((X_{T+t})_{t \geq 0}) \mid \mathcal{F}_T] = \mathbf{1}_{\{T < \infty\}} \mathbf{E}_{X_T}[f((X_t)_{t \geq 0})].$$

Remarque. On dit qu'on utilise la propriété de Markov simple quand on applique la propriété de Markov forte avec un temps déterministe (tout temps déterministe est un temps d'arrêt).

Preuve. Il suffit (admis, voir les compléments en annexe A) de montrer le théorème quand $f((x_t)_{t \geq 0}) = \phi(x_{t_1}, \dots, x_{t_n})$ avec $\phi : E^n \mapsto \mathbb{R}^+$ et $0 \leq t_1 < t_2 < \dots < t_n$. Et on peut supposer que $\phi(x_1, \dots, x_n) = \mathbf{1}_{\{x_1 = i_1, \dots, x_n = i_n\}}$, car toute fonction ϕ sur E^n est une combinaison linéaire de telles fonctions puisque E est dénombrable. L'objectif est donc de montrer que

$$\mathbf{1}_{\{T < \infty\}} \mathbf{P}_v(X_{T+t_1} = i_1, \dots, X_{T+t_n} = i_n \mid \mathcal{F}_T) = \mathbf{1}_{\{T < \infty\}} \mathbf{P}_{X_T}(X_{t_1} = i_1, \dots, X_{t_n} = i_n).$$

Etape 1. On établit d'abord la formule lorsque le temps d'arrêt T est à valeurs dans l'ensemble dénombrable $\{k2^{-\ell} : k \in \mathbb{N}\} \cup \{\infty\}$. Pour cela, on écrit

$$\begin{aligned} I &:= \mathbf{1}_{\{T < \infty\}} \mathbf{P}_v(X_{T+t_1} = i_1, \dots, X_{T+t_n} = i_n \mid \mathcal{F}_T) \\ &= \sum_{k \geq 0} \mathbf{1}_{\{T = k2^{-\ell}\}} \mathbf{P}_v(X_{k2^{-\ell}+t_1} = i_1, \dots, X_{k2^{-\ell}+t_n} = i_n \mid \mathcal{F}_T) \\ &= \sum_{k \geq 0} \mathbf{1}_{\{T = k2^{-\ell}\}} \mathbf{P}_v(X_{k2^{-\ell}+t_1} = i_1, \dots, X_{k2^{-\ell}+t_n} = i_n \mid \mathcal{F}_{k2^{-\ell}}). \end{aligned}$$

Pour la dernière ligne, on a utilisé un argument du type du lemme 4.6. Ceci conduit à

$$\begin{aligned} I &= \sum_{k \geq 0} \mathbf{1}_{\{T = k2^{-\ell}\}} P_{t_1}(X_{k2^{-\ell}}, i_1) P_{t_2-t_1}(i_1, i_2) \cdots P_{t_n-t_{n-1}}(i_{n-1}, i_n) \\ &= \sum_{k \geq 0} \mathbf{1}_{\{T = k2^{-\ell}\}} \mathbf{P}_{X_{k2^{-\ell}}}(X_{t_1} = i_1, \dots, X_{t_n} = i_n) \\ &= \mathbf{1}_{\{T < \infty\}} \mathbf{P}_{X_T}(X_{t_1} = i_1, \dots, X_{t_n} = i_n). \end{aligned}$$

Etape 2. Si maintenant T est un temps d'arrêt quelconque, on l'approche par une suite de temps d'arrêt T_ℓ de sorte que $\lim_\ell \downarrow T_\ell = T$, que pour chaque ℓ , T_ℓ soit à valeurs dans $\{k2^{-\ell} : k \in \mathbb{N}\} \cup \{\infty\}$, et que $\{T_\ell < \infty\} = \{T < \infty\}$ (lemme 2.1). Par l'étape 1, on a donc, pour tout ℓ ,

$$\mathbf{1}_{\{T < \infty\}} \mathbf{P}_v(X_{T_\ell+t_1} = i_1, \dots, X_{T_\ell+t_n} = i_n \mid \mathcal{F}_{T_\ell}) = \mathbf{1}_{\{T < \infty\}} \mathbf{P}_{X_{T_\ell}}(X_{t_1} = i_1, \dots, X_{t_n} = i_n).$$

Comme X est continu à droite et comme T_ℓ décroît vers T , on a $\lim_\ell X_{T_\ell+t} = X_{T+t}$ pour tout t . Donc, en utilisant le fait que $\mathcal{F}_T \subset \mathcal{F}_{T_\ell}$ (et la convergence dominée),

$$\begin{aligned}
& \mathbf{1}_{\{T < \infty\}} \mathbf{P}_V(X_{T+t_1} = i_1, \dots, X_{T+t_n} = i_n \mid \mathcal{F}_T) \\
&= \lim_{\ell \rightarrow \infty} \mathbf{1}_{\{T < \infty\}} \mathbf{P}_V(X_{T_\ell+t_1} = i_1, \dots, X_{T_\ell+t_n} = i_n \mid \mathcal{F}_T) \\
&= \lim_{\ell \rightarrow \infty} \mathbf{1}_{\{T < \infty\}} \mathbf{E}_V[\mathbf{P}_V(X_{T_\ell+t_1} = i_1, \dots, X_{T_\ell+t_n} = i_n \mid \mathcal{F}_{T_\ell}) \mid \mathcal{F}_T] \\
&= \lim_{\ell \rightarrow \infty} \mathbf{E}_V[\mathbf{1}_{\{T < \infty\}} \mathbf{P}_{X_{T_\ell}}(X_{t_1} = i_1, \dots, X_{t_n} = i_n) \mid \mathcal{F}_T] \\
&= \lim_{\ell \rightarrow \infty} \mathbf{E}_V[\mathbf{1}_{\{T < \infty\}} \mathbf{P}_{X_T}(X_{t_1} = i_1, \dots, X_{t_n} = i_n) \mid \mathcal{F}_T] \\
&= \mathbf{E}_V[\mathbf{1}_{\{T < \infty\}} \mathbf{P}_{X_T}(X_{t_1} = i_1, \dots, X_{t_n} = i_n)] \\
&= \mathbf{1}_{\{T < \infty\}} \mathbf{P}_{X_T}(X_{t_1} = i_1, \dots, X_{t_n} = i_n).
\end{aligned}$$

Pour l'avant dernière inégalité, on a utilisé que, comme X est continu à droite pour la distance discrète et comme T_ℓ décroît vers T , p.s., $X_{T_\ell} = X_T$ pour ℓ assez grand (cf Remarque ci-après). $\square$

5.2 Description dynamique

Considérons un PMS $(X_t)_{t \geq 0}$ et posons $T_0 = 0$ puis, par récurrence, pour $n \geq 0$,

$$T_{n+1} = \inf\{t > T_n : X_t \neq X_{T_n}\}.$$

Les T_n , $n \geq 0$, sont des temps d'arrêt. De plus, si on suppose que pour tout $i \in E$, $\mathbf{P}_i(T_1 < \infty) = 1$, alors p.s., pour tout $n \geq 1$, $T_n < \infty$. Cela se montre par récurrence, en utilisant Markov forte. Par exemple, en posant $\tilde{X}_t = X_{T_1+t}$, on a $\tilde{T}_1 = T_2 - T_1$, et donc

$$\begin{aligned}
\mathbf{P}_i(T_2 < \infty) &= \mathbf{E}_i[\mathbf{1}_{T_1 < \infty} \mathbf{P}_i(T_2 - T_1 < \infty \mid \mathcal{F}_{T_1})] \\
&= \mathbf{E}_i[\mathbf{1}_{T_1 < \infty} \mathbf{P}_i(\tilde{T}_1 < \infty \mid \mathcal{F}_{T_1})] \\
&= \mathbf{E}_i[\mathbf{1}_{T_1 < \infty} \mathbf{P}_{X_{T_1}}(T_1 < \infty)] \\
&= \mathbf{E}_i[\mathbf{1}_{T_1 < \infty}] \\
&= \mathbf{P}_i(T_1 < \infty) = 1.
\end{aligned}$$

On rappelle la propriété suivante des loi exponentielles: si $U \sim \text{Exp}(1)$ et $\lambda > 0$, alors $U/\lambda \sim \text{Exp}(\lambda)$.

Théorème 5.4. On suppose que pour tout $i \in E$, $\mathbf{P}_i(T_1 < \infty) = 1$, et on pose $Q(i, j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i(T_1)$, pour $i, j \in E$.

- (i) $(X_{T_n})_{n \in \mathbb{N}}$ est une chaîne de Markov de transition Q .
- (ii) Sachant $(X_{T_n})_{n \in \mathbb{N}}$, les v.a. $T_1, T_2 - T_1, \dots$ sont indépendantes de lois $\text{Exp}(\lambda(X_{T_0})), \text{Exp}(\lambda(X_{T_1})), \dots$

Les $(\lambda(i))_{i \in E}$ sont les “taux de saut” du PMS et $(Q(i, j))_{i, j \in E}$ est sa “matrice de transition”. Voici une autre façon de dire la même chose, qui va nous conduire à une procédure pour simuler les processus markovien de sauts.

Corollaire 5.5. Avec les mêmes hypothèses et notations que dans le théorème, il existe une chaîne de Markov $(\xi_n)_{n \geq 0}$ de transition Q , indépendante d'une suite $(U_n)_{n \geq 1}$ de v.a. indépendantes de loi exponentielle de paramètre 1 telles que, pour tout $t \in \mathbb{R}^+$,

$$X_t = \sum_{n=0}^{+\infty} \xi_n \mathbf{1}_{\{t \in [T_n, T_{n+1}[\}}$$

où $T_0 = 0$ et $T_{n+1} = T_n + U_{n+1}/\lambda(\xi_n)$.

Preuve. Il suffit de poser $\xi_n = X_{T_n}$ et $U_n = \lambda(\xi_{n-1})(T_n - T_{n-1})$ et d'appliquer le théorème. $\square$

Preuve du théorème. Etape 1. Montrons déjà que les temps d'attente sont exponentiels: pour tout $i \in E$, $\mathbf{P}_i(T_1 \geq t) = e^{-\lambda(i)t}$ pour tout $t \geq 0$. On écrit

$$\begin{aligned} \mathbf{P}_i(T_1 > t + s) &= \mathbf{P}_i(\forall u \in [0, t + s], X_u = i) \\ &= \mathbf{E}_i[\mathbf{1}_{\{\forall u \in [0, t], X_u = i\}} \mathbf{P}_i(\forall u \in [t, t + s], X_u = i \mid \mathcal{F}_t)]. \end{aligned}$$

Mais par la propriété de Markov simple (on introduit $\tilde{X}_s = X_{t+s}$),

$$\begin{aligned} \mathbf{P}_i(\forall u \in [t, t + s], X_u = i \mid \mathcal{F}_t) &= \mathbf{P}_i(\forall u \in [0, s], \tilde{X}_u = i \mid \mathcal{F}_t) \\ &= \mathbf{P}_{X_t}(\forall u \in [0, s], X_u = i). \end{aligned}$$

Ainsi,

$$\begin{aligned} \mathbf{P}_i(T_1 > t + s) &= \mathbf{E}_i[\mathbf{1}_{\{\forall u \in [0, t], X_u = i\}} \mathbf{P}_i(\forall u \in [0, s], X_u = i)] \\ &= \mathbf{P}_i(\forall u \in [0, t], X_u = i) \times \mathbf{P}_i(\forall u \in [0, s], X_u = i) \\ &= \mathbf{P}_i(T_1 > t) \times \mathbf{P}_i(T_1 > s). \end{aligned}$$

Donc T_1 est sans-mémoire et suit une loi exponentielle sous $\mathbf{P}_i$. Son paramètre est bien sûr $\lambda(i) = 1/\mathbf{E}_i[T_1]$.

Etape 2. Remarquons ensuite que pour tout $i, j \in E$, et $t \geq 0$,

$$\mathbf{P}_i(T_1 > t, X_{T_1} = j) = e^{-\lambda(i)t} Q(i, j).$$

En effet, par Markov simple (avec $\tilde{X}_s = X_{t+s}$, on a $\tilde{X}_{\tilde{T}_1} = X_{T_1}$ sur $\{T_1 > t\}$),

$$\begin{aligned} \mathbf{P}_i(T_1 > t, X_{T_1} = j) &= \mathbf{E}_i[\mathbf{1}_{\{T_1 > t\}} \mathbf{P}_i(X_{T_1} = j \mid \mathcal{F}_t)] \\ &= \mathbf{E}_i[\mathbf{1}_{\{T_1 > t\}} \mathbf{P}_i(\tilde{X}_{\tilde{T}_1} = j \mid \mathcal{F}_t)] \\ &= \mathbf{E}_i[\mathbf{1}_{\{T_1 > t\}} \mathbf{P}_{X_t}(X_{T_1} = j)]. \end{aligned}$$

Puis, comme $T_1 > t$ implique que $X_t = i$,

$$\mathbf{P}_i(T_1 > t, X_{T_1} = j) = \mathbf{E}_i[\mathbf{1}_{\{T_1 > t\}} \mathbf{P}_i(X_{T_1} = j)] = \mathbf{E}_i[\mathbf{1}_{\{T_1 > t\}} Q(i, j)]$$

par définition de Q , et on conclut à l'aide de l'étape 1.

Etape 3. Pour $n \geq 1$, $u_1, \dots, u_n > 0$ et $i_1, \dots, i_n \in E$,

$$\begin{aligned} \mathbf{P}_i(T_1 > u_1, X_{T_1} = i_1, \dots, T_n - T_{n-1} > u_n, X_{T_n} = i_n) \\ = \mathbf{E}_i[\mathbf{1}_{\{T_1 > u_1, X_{T_1} = i_1, \dots, T_{n-1} - T_{n-2} > u_n, X_{T_{n-1}} = i_{n-1}\}} \mathbf{E}_i(\mathbf{1}_{\{T_n - T_{n-1} > u_n, X_{T_n} = i_n\}} \mid \mathcal{F}_{T_{n-1}})]]. \end{aligned}$$

Par la propriété de Markov forte en T_{n-1} (avec $\tilde{X}_t = X_{T_{n-1}+t}$, noter que $\tilde{T}_1 = T_n - T_{n-1}$ et que $\tilde{X}_{\tilde{T}_1} = X_{T_n}$),

$$\begin{aligned} \mathbf{E}_i(\mathbf{1}_{\{T_n - T_{n-1} > u_n, X_{T_n} = i_n\}} \mid \mathcal{F}_{T_{n-1}}) &= \mathbf{E}_i(\mathbf{1}_{\{\tilde{T}_1 > u_n, \tilde{X}_{\tilde{T}_1} = i_n\}} \mid \mathcal{F}_{T_{n-1}}) \\ &= \mathbf{E}_{X_{T_{n-1}}}(\mathbf{1}_{\{T_1 > u_n, X_{T_1} = i_n\}}), \end{aligned}$$

qui vaut $Q(X_{T_{n-1}}, i_n)e^{-\lambda(X_{T_{n-1}})u_n}$ par l'étape 2. Ainsi,

$$\begin{aligned} \mathbf{P}_i(T_1 > u_1, X_{T_1} = i_1, \dots, T_n - T_{n-1} > u_n, X_{T_n} = i_n) \\ &= \mathbf{E}_i[\mathbf{1}_{\{T_1 > u_1, X_{T_1} = i_1, \dots, T_{n-1} - T_{n-2} > u_{n-1}, X_{T_{n-1}} = i_{n-1}\}} Q(X_{T_{n-1}}, i_n)e^{-\lambda(X_{T_{n-1}})u_n}] \\ &= \mathbf{P}_i(T_1 > u_1, X_{T_1} = i_1, \dots, T_{n-1} - T_{n-2} > u_{n-1}, X_{T_{n-1}} = i_{n-1}) \times Q(i_{n-1}, i_n)e^{-\lambda(i_{n-1})u_n}. \end{aligned}$$

On montre donc, en répétant ce procédé, que

$$\begin{aligned} \mathbf{P}_i(T_1 > u_1, X_{T_1} = i_1, \dots, T_n - T_{n-1} > u_n, X_{T_n} = i_n) \\ &= Q(i, i_1)e^{-\lambda(i)u_1} Q(i_1, i_2)e^{-\lambda(i_1)u_2} \dots Q(i_{n-1}, i_n)e^{-\lambda(i_{n-1})u_n}. \end{aligned}$$

En choisissant $u_1 = \dots = u_n = 0$, on conclut aisément que $(X_{T_n})_{n \geq 0}$ est une chaîne de Markov de transition Q , puisque

$$\mathbf{P}_i(X_{T_1} = i_1, \dots, X_{T_n} = i_n) = Q(i, i_1)Q(i_1, i_2) \dots Q(i_{n-1}, i_n).$$

Enfin, on voit que

$$\begin{aligned} \mathbf{P}_i(T_1 > u_1, \dots, T_n - T_{n-1} > u_n \mid X_{T_1} = i_1, \dots, X_{T_n} = i_n) \\ &= \frac{\mathbf{P}_i(T_1 > u_1, X_{T_1} = i_1, \dots, T_n - T_{n-1} > u_n, X_{T_n} = i_n)}{\mathbf{P}_i(X_{T_1} = i_1, \dots, X_{T_n} = i_n)} \\ &= e^{-\lambda(i)u_1} e^{-\lambda(i_1)u_2} \dots e^{-\lambda(i_{n-1})u_n}. \end{aligned}$$

Ceci achève la preuve. □

Au cours de la preuve du théorème, on a en particulier montré

Proposition 5.6. Pour tout $i, j \in E$, $\mathbf{P}_i(X_{T_1} = j, T_1 > t) = e^{-\lambda(i)t} Q(i, j)$.

Algorithme (Simulation d'un PMS). On se donne une matrice de transition $(Q(i, j))_{i, j \in E}$, des taux de saut $(\lambda(i))_{i \in E}$, et une loi initiale ν sur E .

Initialisation:

- on simule l'état initial ξ_0 de loi ν ,
- on pose $T_0 = 0$.

Itérations: supposons qu'on a simulé (ξ_i, T_i) pour $1 \leq i \leq n$ jusqu'à un certain $n \geq 0$, et défini le processus X_t sur l'intervalle $[0, T_n]$, alors

- on simule $U_{n+1} \sim \text{Exp}(1)$ et on pose $T_{n+1} = T_n + U_{n+1} / \lambda(\xi_n)$,
- pour $t \in [T_n, T_{n+1}[$ on pose $X_t = \xi_n$.
- on simule $\xi_{n+1} \sim Q(\xi_n, \cdot)$

Enfin, une remarque qui peut s'avérer pratique.

Proposition 5.7. Soit $(Q(i, j))_{i, j \in E}$ une matrice de transition t.q. $Q(i, i) = 0$ pour tout $i \in E$ et soit $(\lambda(i))_{i \in E}$ une famille de réels positifs tels que $\sup_{i \in E} \lambda(i) \leq \lambda < \infty$. Soit $(Z_n)_{n \geq 0}$ une chaîne de Markov de transition P définie par

$$P(i, i) = 1 - \frac{\lambda(i)}{\lambda} \quad \text{et} \quad P(i, j) = \frac{\lambda(i)Q(i, j)}{\lambda} \quad \text{si } i \neq j.$$

Soit aussi $(N_t)_{t \geq 0}$ un processus de Poisson de paramètre λ , indépendant de $(Z_n)_{n \geq 0}$. Alors $(Z_{N_t})_{t \geq 0}$ est un PMS de taux de sauts $(\lambda(i))_{i \in E}$ et de matrice de transition $(Q(i, j))_{i, j \in E}$.

Preuve. On sait déjà que $(Z_{N_t})_{t \geq 0}$ est un PMS (voir le lemme 5.1). Notons T_1 l'instant du premier saut de $(Z_{N_t})_{t \geq 0}$. On doit montrer que $\mathbf{E}_i[T_1] = 1/\lambda(i)$ et que $\mathbf{P}_i(Z_{N_{T_1}} = j) = Q(i, j)$.

$$\mathbf{P}_i(T_1 > t) = \sum_{k \geq 0} \mathbf{P}_i(N_t = k, Z_1 = \dots = Z_k = i) = \sum_{k \geq 0} e^{-\lambda t} \frac{(\lambda t)^k}{k!} (P(i, i))^k,$$

d'où $\mathbf{P}_i(T_1 > t) = \exp(-\lambda(1 - P(i, i))t) = \exp(-\lambda(i)t)$. Donc $\mathbf{E}_i[T_1] = 1/\lambda(i)$. Ensuite, pour $j \neq i$, (si $i = j$, on a $Q(i, i) = 0 = \mathbf{P}_i[Z_{N_{T_1}} = i]$)

$$\begin{aligned} \mathbf{P}_i(Z_{N_{T_1}} = j) &= \sum_{\ell \geq 1} \mathbf{P}_i(Z_1 = \dots = Z_{\ell-1} = i, Z_\ell = j) \\ &= \sum_{\ell \geq 1} (P(i, i))^{\ell-1} P(i, j) \\ &= \frac{P(i, j)}{1 - P(i, i)} \\ &= Q(i, j). \end{aligned}$$

□

Remarque. On peut voir le résultat suivant rapidement de la manière suivante:

- (i) Soient $\tau_1 < \tau_2 < \dots$ les temps de saut du processus de Poisson $(N_t)_{t \geq 0}$. Comme $(Z_n)_{n \geq 0}$ est indépendante de N , sous $\mathbf{P}_i$, T_1 est le premier des τ_n que l'on garde quand on efface ceux pour lesquels $Z_n = i$. Tant qu'on a pas quitté l'état i , on garde un point avec probabilité $\lambda(i)/\lambda$, indépendamment de N ; donc $T_1 \sim \text{Exp}(\lambda(i))$ (processus de Poisson marqué).
- (ii) De plus, par l'argument précédent, en T_1 la loi de $Z_{N_{T_1}}$ sous $\mathbf{P}_i$ est précisément $P(i, \cdot)$ conditionnée à être différent de i , d'où

$$\mathbf{P}_i(Z_{N_{T_1}} = j) = \frac{P(i, j)}{1 - P(i, i)} = Q(i, j).$$

5.3 Explosion

Définition (Explosion). On dit qu'il y a **explosion** pour un processus markovien de saut si, avec probabilité positive, il y a un nombre infini de sauts dans un intervalle de temps fini.

Pour être plus concret, on se donne un espace d'états E dénombrable, des taux de saut $(\lambda(i))_{i \in E}$ une matrice de transition $(Q(i, j))_{i, j \in E}$, et on construit le PMS en utilisant le procédé de simulation plus haut. Il est a priori possible que, avec une probabilité non nulle, $\lim_{n \rightarrow +\infty} T_n = \zeta < +\infty$. Il y a alors explosion, et

il n'est alors pas clair comment définir le processus au delà du temps ζ . C'est assez ennuyeux, puisqu'il est alors impossible de parler de X_t , sans être sûr qu'il n'y a pas eu explosion. Les critères garantissant que le processus n'explose pas sont donc particulièrement importants.

En vue de donner un critère de non explosion, montrons d'abord:

Lemme 5.8. Soit $U_n, n \geq 1$, une suite de variables aléatoires indépendantes de lois exponentielles de paramètre $\lambda_1, \lambda_2, \dots$.

- (i) Si $\sum_{k=1}^{\infty} 1/\lambda_k < \infty$, alors $\sum_{k=1}^{\infty} U_k < \infty$ p.s.
- (ii) Si $\sum_{k=1}^{\infty} 1/\lambda_k = \infty$, alors $\sum_{k=1}^{\infty} U_k = \infty$ p.s.

Preuve. Comme $\mathbf{E}[\sum_{k=1}^{\infty} U_k] = \sum_{k=1}^{+\infty} 1/\lambda_k$, le point (i) est trivial. Pour (ii), en utilisant l'indépendance, on écrit

$$\mathbf{E}\left[\exp\left(-\sum_{k=1}^{\infty} U_k\right)\right] = \prod_{k=1}^{\infty} \mathbf{E}[e^{-U_k}].$$

Mais si $U \sim \text{Exp}(1)$, alors $V = e^{-U}$ est uniforme sur $[0, 1]$. Par conséquent, comme $U_k \sim \text{Exp}(\lambda_k)$ a la même distribution que U/λ_k on a $\mathbf{E}[e^{-U_k}] = \mathbf{E}[V^{1/\lambda_k}]$, ce qui implique

$$\mathbf{E}\left[\exp\left(-\sum_{k=1}^{\infty} U_k\right)\right] = \prod_{k=1}^{\infty} \frac{\lambda_k}{1 + \lambda_k} = \exp\left(-\sum_{k=1}^{\infty} \log(1 + 1/\lambda_k)\right),$$

Il y a maintenant deux possibilités:

- soit $\lambda_k \not\rightarrow \infty$ auquel cas $\log(1 + 1/\lambda_k) \not\rightarrow 0$ et la série diverge;
- soit $\lambda_k \rightarrow \infty$ auquel cas $\log(1 + 1/\lambda_k) \geq 1/(2\lambda_k)$ pour k assez grand, et la série diverge par hypothèse.

Par conséquent, indépendamment de la situation, on a $\mathbf{E}[\exp(-\sum_{k=1}^{\infty} U_k)] = 0$, ce qui implique que $\sum_{k=1}^{\infty} U_k = \infty$ p.s. $\square$

Proposition 5.9 (Un critère de non explosion). Considérons un PMS $(X_t)_{t \geq 0}$. On a

$$\mathbf{P}\left(\lim_{n \rightarrow +\infty} T_n < \infty\right) = \mathbf{P}\left(\sum_{n=0}^{\infty} \frac{1}{\lambda(X_{T_n})} < \infty\right).$$

En particulier, si $\sum_{n=0}^{+\infty} \frac{1}{\lambda(X_{T_n})} = +\infty$ p.s., alors il n'y a pas explosion.

Preuve. En utilisant le lemme ci-dessus et le fait que sachant $(X_{T_n})_{n \geq 0}$, les $T_{n+1} - T_n$ sont indépendants et $\text{Exp}(\lambda(X_{T_n}))$, on peut écrire que

$$\begin{aligned} \mathbf{P}\left(\lim_{n \rightarrow +\infty} T_n < \infty\right) &= \mathbf{P}\left(\sum_{n \geq 0} (T_{n+1} - T_n) < \infty\right) \\ &= \mathbf{E}\left[\mathbf{P}\left(\sum_{n \geq 0} (T_{n+1} - T_n) < \infty \mid (X_{T_n})_{n \geq 0}\right)\right] \\ &= \mathbf{E}\left[\mathbf{1}_{\{\sum_{n=0}^{\infty} \frac{1}{\lambda(X_{T_n})} < \infty\}}\right] \\ &= \mathbf{P}\left(\sum_{n=0}^{\infty} \frac{1}{\lambda(X_{T_n})} < \infty\right). \end{aligned}$$

$\square$

Corollaire 5.10. Soit un PMS $(X_t)_{t \geq 0}$. Si l'une des conditions suivantes est satisfaite

- (i) $\sup_{i \in E} \lambda(i) < \infty$,
 - (ii) la chaîne $(X_{T_n})_{n \geq 0}$ est récurrente,
- alors il n'y a pas d'explosion.

Preuve. Il résulte de la proposition précédente que, pour qu'il y ait explosion, il faut que $\lambda(X_{T_n}) \rightarrow +\infty$ lorsque $n \rightarrow +\infty$ avec probabilité positive. C'est impossible dans les deux cas (i) et (ii). $\square$

Remarque. L'explosion entraîne de nombreux phénomènes désagréables. Pour les éviter **nous supposons toujours désormais qu'il n'y a pas d'explosion.**

5.4 Le générateur

Quand on modélise un phénomène, on part généralement des taux d'évènement $(\lambda(i))_{i \in E}$ et de la matrice de transition $(Q(i, j))_{i, j \in E}$. S'il est presque toujours impossible de décrire explicitement le semigroupe $(P_t)_{t \geq 0}$, on peut souvent l'étudier, en utilisant les équations de Kolmogorov. Commençons par définir le générateur, qui n'est pour l'instant qu'une notation.

Définition (Générateur). Le générateur du PMS est la "matrice" $(A(i, j))_{i, j \in E}$ définie par

$$A(i, j) = \begin{cases} \lambda(i)Q(i, j), & \text{si } i \neq j, \\ -\lambda(i), & \text{si } i = j. \end{cases}$$

Remarque. Comme $Q(i, i) = 0$, on a $\sum_{j \neq i} Q(i, j) = 1$ et donc les sommes par ligne de A sont toutes nulles: $\sum_{j \in E} A(i, j) = 0$.

Pour décrire le lien entre le générateur et le semigroupe, montrons le lemme important suivant qui dit qu'en un temps petit, le processus n'a pas le temps de sauter deux fois.

Lemme 5.11. Pour tout $i \in E$, $\mathbf{P}_i(T_2 \leq t) = o(t)$ quand $t \rightarrow 0$.

Preuve. On écrit

$$\mathbf{P}_i(T_2 \leq t) = \sum_{j \in E} \mathbf{P}_i(X_{T_1} = j, T_2 \leq t) = \sum_{j \in E} \mathbf{P}_i(X_{T_1} = j) \mathbf{P}_i(T_2 \leq t \mid X_{T_1} = j).$$

Mais par le théorème de la section précédente, sachant $X_{T_1} = j$, les v.a. T_1 et $T_2 - T_1$ sont indépendantes et respectivement de lois $\text{Exp}(\lambda(i))$ et $\text{Exp}(\lambda(j))$, d'où

$$\begin{aligned} \mathbf{P}_i(T_2 \leq t \mid X_{T_1} = j) &\leq \mathbf{P}_i(T_1 \leq t, T_2 - T_1 \leq t \mid X_{T_1} = j) \\ &= (1 - e^{-\lambda(i)t})(1 - e^{-\lambda(j)t}) \\ &\leq \lambda(i)t(1 - e^{-\lambda(j)t}). \end{aligned}$$

Donc

$$\mathbf{P}_i(T_2 \leq t) \leq \lambda(i)t \sum_{j \in E} \mathbf{P}_i(X_{T_1} = j)(1 - e^{-\lambda(j)t}).$$

C'est un $o(t)$ car $\lim_{t \rightarrow 0} \sum_{j \in E} \mathbf{P}_i(X_{T_1} = j)(1 - e^{-\lambda(j)t}) = 0$ par convergence dominée. $\square$

Théorème 5.12 (Équations de Kolmogorov). Le semigroupe $(P_t)_{t \geq 0}$ vérifie $P_0 = I$ et, pour tout $t \geq 0$,

- *équation de Kolmogorov backward*:

$$\frac{dP_t}{dt} = AP_t;$$

- *équation de Kolmogorov forward*: s'il n'y a pas explosion,

$$\frac{dP_t}{dt} = P_t A.$$

Remarque. L'équation de Kolmogorov backward se réécrit : pour tout $i, j \in E$,

$$P'_t(i, j) = -\lambda(i)P_t(i, j) + \sum_{k \neq i} \lambda(i)Q(i, k)P_t(k, j).$$

L'équation de Kolmogorov forward se réécrit : pour tout $i, j \in E$,

$$P'_t(i, j) = -P_t(i, j)\lambda(j) + \sum_{k \neq j} P_t(i, k)\lambda(k)Q(k, j).$$

Nous allons montrer proprement l'équation backward et "rapidement" l'équation forward.

Preuve rapide de l'équation forward. (Cette preuve est correcte si E est fini, l'explosion étant alors impossible). Pour $0 \leq t < t+h$, on écrit

$$P_{t+h}(i, j) = \mathbf{P}_i(X_{t+h} = j) = I + J + K,$$

où

$$I = \mathbf{P}_i(X_{t+h} = j \text{ et pas de saut entre } t \text{ et } t+h),$$

$$J = \mathbf{P}_i(X_{t+h} = j \text{ et exactement un saut entre } t \text{ et } t+h),$$

$$K = \mathbf{P}_i(X_{t+h} = j \text{ et au moins deux sauts entre } t \text{ et } t+h).$$

Par la propriété de Markov simple en t , en posant $\tilde{X}_s = X_{t+s}$, on a

$$\begin{aligned} I &= \mathbf{P}_i(X_t = j \text{ et pas de saut entre } t \text{ et } t+h), \\ &= \mathbf{E}_i[\mathbf{1}_{\{X_t=j\}} \mathbf{P}_i(\tilde{T}_1 > h \mid \mathcal{F}_t)] \\ &= \mathbf{P}_i(X_t = j) \mathbf{P}_j(T_1 > h) \\ &= P_t(i, j) e^{-\lambda(j)h} \\ &= P_t(i, j)(1 - \lambda(j)h) + o(h). \end{aligned}$$

Toujours par la propriété de Markov simple appliquée en t (et par le lemme), en posant encore $\tilde{X}_s = X_{t+s}$,

$$\begin{aligned} J &= \sum_{k \neq j} \mathbf{P}_i(X_t = k \text{ et exactement un saut, de } k \text{ à } j, \text{ entre } t \text{ et } t+h) \\ &= \sum_{k \neq j} \mathbf{E}_i[\mathbf{1}_{\{X_t=k\}} \mathbf{P}_i(\tilde{T}_1 < h < \tilde{T}_2, \tilde{X}_{\tilde{T}_1} = j \mid \mathcal{F}_t)] \\ &= \sum_{k \neq j} \mathbf{P}_i(X_t = k) \mathbf{P}_k(T_1 < h < T_2, X_{T_1} = j). \end{aligned}$$

Mais $\mathbf{P}_k(T_1 < h < T_2, X_{T_1} = j) = \mathbf{P}_k(T_1 < h, X_{T_1} = j) - \mathbf{P}_k(T_2 < h, X_{T_1} = j) = \mathbf{P}_k(T_1 < h, X_{T_1} = j) + o(h)$ par le lemme, d'où (voir la remarque plus bas pour le passage de $(\star)$ à $(\star\star)$)

$$\begin{aligned} J &= \sum_{k \neq j} \mathbf{P}_i(X_t = k) \cdot [\mathbf{P}_k(T_1 < h, X_{T_1} = j) + o(h)] \quad (\star) \\ &= \sum_{k \neq j} P_t(i, k)(1 - e^{-\lambda(k)h})Q(k, j) + o(h) \quad (\star\star) \\ &= \sum_{k \neq j} P_t(i, k)\lambda(k)Q(k, j)h + o(h). \end{aligned}$$

Enfin, par la propriété de Markov simple, avec toujours $\tilde{X}_s = X_{t+s}$,

$$K = \mathbf{E}_i[\mathbf{P}_i(\tilde{X}_h = j, \tilde{T}_2 \leq h \mid \mathcal{F}_t)] \leq \mathbf{E}_i[\mathbf{P}_i(\tilde{T}_2 \leq h \mid \mathcal{F}_t)] = \mathbf{E}_i[\mathbf{P}_{X_t}(T_2 \leq h)] = o(h) \quad (\star\star\star)$$

par le lemme. Ainsi,

$$P_{t+h}(i, j) = P_t(i, j)(1 - \lambda(j)h) + \sum_{k \neq j} P_t(i, k)\lambda(k)Q(k, j)h + o(h),$$

donc

$$\frac{P_{t+h}(i, j) - P_t(i, j)}{h} = -P_t(i, j)\lambda(j) + \sum_{k \neq j} P_t(i, k)\lambda(k)Q(k, j) + \frac{o(h)}{h},$$

et il n'y a plus qu'à faire tendre h vers 0. □

Remarque. Dans la preuve précédente, le problème lorsque E est infini est le passage de $(\star)$ à $(\star\star)$, ainsi que fait que $(\star\star\star)$ est effectivement un $o(h)$. En effet, le $o(h)$ en $(\star)$ dépend a priori de $k \in E$, et si on l'explicite comme $\varepsilon_k(h)$ alors ce qu'on prétend en $(\star\star)$, c'est que

$$\sum_{k \neq j} \mathbf{P}_i(X_t = k)\varepsilon_k(h) = o(h),$$

ce qui n'est pas clair. D'autant plus que $\varepsilon_k(h) = \mathbf{P}_k(T_2 < h, X_{T_1} = j)$ dépend de λ_k puisque sous $\mathbf{P}_k$, on a $T_1 = E/\lambda_k$ avec $E \sim \text{Exp}(1)$; $\mathbf{P}_k(T_2 < h)$ est donc d'autant plus grande à h fixé que λ_k est grand. Le même problème intervient en $(\star\star\star)$. Ceci explique en partie pourquoi l'équation forward peut ne pas être vraie s'il y a explosion.

Preuve de l'équation backward. Etape 1. On montre d'abord que pour tout $i, j \in E$, tout $t \geq 0$, on a l'expression suivante

$$P_t(i, j) = e^{-\lambda(i)t} \mathbf{1}_{\{i=j\}} + \sum_{k \neq i} \int_0^t e^{-\lambda(i)(t-s)} A(i, k) P_s(k, j) ds. \quad (\star)$$

Pour cela, on écrit

$$\begin{aligned}
 P_t(i, j) &= \mathbf{P}_i(X_t = j) \\
 &= \mathbf{P}_i(X_t = j, T_1 > t) + \mathbf{P}_i(X_t = j, T_1 \leq t) \\
 &= \mathbf{1}_{\{i=j\}} e^{-\lambda(i)t} + \sum_{k \neq i} \mathbf{P}_i(X_{T_1} = k, X_t = j, T_1 \leq t).
 \end{aligned}$$

Mais par la propriété de Markov forte en T_1 , en posant $\tilde{X}_t = X_{T_1+t}$, on obtient

$$\begin{aligned}
 \mathbf{P}_i(X_{T_1} = k, X_t = j, T_1 \leq t) &= \mathbf{E}_i[\mathbf{1}_{\{X_{T_1}=k, T_1 \leq t\}} \mathbf{P}_i(X_t = j \mid \mathcal{F}_{T_1})] \\
 &= \mathbf{E}_i[\mathbf{1}_{\{X_{T_1}=k, T_1 \leq t\}} \mathbf{P}_i(\tilde{X}_{t-T_1} = j \mid \mathcal{F}_{T_1})] \\
 &= \mathbf{E}_i[\mathbf{1}_{\{X_{T_1}=k, T_1 \leq t\}} P_{t-T_1}(X_{T_1}, j)] \\
 &= \mathbf{E}_i[\mathbf{1}_{\{X_{T_1}=k, T_1 \leq t\}} P_{t-T_1}(k, j)].
 \end{aligned}$$

En se rappelant que $\mathbf{P}_i(T_1 > t, X_{T_1} = j) = e^{-\lambda(i)t} Q(i, j)$ (Proposition 5.6), on voit que T_1 et X_{T_1} sont indépendants, que $T_1 \sim \text{Exp}(\lambda(i))$ (sous $\mathbf{P}_i$) et que $\mathbf{P}_i(X_{T_1} = k) = Q(i, k)$. Ainsi,

$$\begin{aligned}
 \mathbf{P}_i(X_{T_1} = k, X_t = j, T_1 \leq t) &= \int_0^t e^{-\lambda(i)s} P_{t-s}(k, j) \lambda(i) Q(i, k) ds \\
 &= \int_0^t e^{-\lambda(i)(t-s)} A(i, k) P_s(k, j) ds.
 \end{aligned}$$

On a finalement utilisé que, comme $k \neq i$, $A(i, k) = \lambda(i) Q(i, k)$, et le changement de variables $s \rightarrow t - s$.

Etape 2. L'expression en (★) se factorise en

$$P_t(i, j) = e^{-\lambda(i)t} \left[\mathbf{1}_{\{i=j\}} + \sum_{k \neq i} \int_0^t e^{\lambda(i)s} A(i, k) P_s(k, j) ds \right].$$

En utilisant le fait que P_s est borné par 1 et comme $\sum_{k \neq i} A(i, k) = \lambda(i)$, on déduit de cette formule (par convergence dominée) que l'application $t \mapsto P_t(i, j)$ est continue, puis dans un second temps qu'elle est dérivable et que sa dérivée est donnée par

$$\begin{aligned}
 P_t'(i, j) &= -\lambda(i) P_t(i, j) + e^{-\lambda(i)t} \left[\sum_{k \neq i} e^{\lambda(i)t} A(i, k) P_t(k, j) \right] \\
 &= -\lambda(i) P_t(i, j) + \sum_{k \neq i} A(i, k) P_t(k, j)
 \end{aligned}$$

comme désiré. □

5.5 Mesures invariantes

Définition. (i) On dit qu'un PMS $(X_t)_{t \geq 0}$ est :

- irréductible si la chaîne $(X_{T_n})_{n \geq 0}$ de transition Q est irréductible,
- récurrent si la chaîne $(X_{T_n})_{n \geq 0}$ de transition Q est récurrente,
- transitoire si la chaîne $(X_{T_n})_{n \geq 0}$ de transition Q est transitoire.

(ii) On dit que m est une mesure invariante si pour tout $t \geq 0$, $m P_t = m$.

(iii) On dit que le PMS est récurrent positif s'il admet une probabilité invariante.

Remarque. (i) La récurrence/transience d'un PMS ne dépend que de sa matrice de transition / chaîne de Markov sous-jacente, alors que le fait d'être récurrent positif est une propriété qui fait intervenir le semi-groupe $(P_t)_{t \geq 0}$. En particulier, il n'est pas immédiatement clair qu'un PMS récurrent positif soit récurrent. La proposition 5.14 plus bas montre que c'est effectivement le cas.
(ii) Nous verrons à la fin de la section qu'il est possible que le PSM $(X_t)_{t \geq 0}$ soit récurrent positif sans que Q le soit (Exemple 5.1), ou inversement (Exemple 5.2).

Lemme 5.13. Si le PMS $(X_t)_{t \geq 0}$ est irréductible et récurrent, la chaîne de Markov de transition

$$\Pi(i, j) = \int_0^\infty e^{-t} P_t(i, j) dt$$

est aussi irréductible et récurrente.

Remarque. (i) Noter que pour tout $i \in E$, on a $\sum_{j \in E} \Pi(i, j) = \int_0^\infty e^{-t} dt = 1$: Π est bien une matrice de transition.

(ii) De plus, si $\tau_0 = 0$ et si $\tau_1 < \tau_2 < \dots$ sont les temps de sauts d'un processus de Poisson d'intensité 1, alors Π est la matrice de transition de la chaîne de Markov $(X_{\tau_n})_{n \geq 0}$, puisque $(\tau_{n+1} - \tau_n)_{n \geq 0}$ sont i.i.d. de loi $\text{Exp}(1)$ et $\Pi(i, j) = \mathbf{E}[P_{\tau_1}(i, j)] = \mathbf{P}_i(X_{\tau_1} = j)$.

Preuve. Par définition, comme $(X_t)_{t \geq 0}$ est récurrent, on a $\sum_{n \geq 0} Q^n(i, j) = \infty$ pour tout $i, j \in E$. Nous allons montrer que

$$\sum_{n \geq 1} \Pi^n(i, j) = \frac{1}{\lambda(j)} \sum_{n \geq 0} Q^n(i, j), \quad (\star)$$

ce qui permettra d'affirmer que $\sum_{n \geq 1} \Pi^n(i, j) = \infty$, et donc que la chaîne de transition Π est irréductible et récurrente.

Pour celà, montrons d'abord par récurrence sur $n \geq 0$ que

$$\Pi^n(i, j) = \int_0^\infty e^{-t} \frac{t^{n-1}}{(n-1)!} P_t(i, j) dt. \quad (\star\star)$$

C'est vrai si $n = 1$ par définition. Supposons maintenant que c'est vrai pour $n \geq 0$. Alors, comme $P_t P_s = P_{t+s}$,

$$\Pi^{n+1} = \left(\int_0^\infty e^{-t} \frac{t^{n-1}}{(n-1)!} P_t dt \right) \left(\int_0^\infty e^{-s} P_s ds \right) = \int_0^\infty \int_0^\infty e^{-t-s} \frac{t^{n-1}}{(n-1)!} P_{t+s} dt ds,$$

puis, en changeant de variables $(s, t) \rightarrow (u = t + s, t)$,

$$\Pi^{n+1} = \int_0^\infty e^{-u} P_u \int_0^u \frac{t^{n-1}}{(n-1)!} dt du = \int_0^\infty e^{-u} P_u \frac{u^n}{n!} du,$$

ce qui prouve que la formule est aussi vraie pour $n + 1$, et termine la preuve de $(\star\star)$.

Revenons maintenant à la preuve de $(\star)$. En utilisant l'expression $(\star\star)$, on obtient

$$\sum_{n \geq 1} \Pi^n(i, j) = \sum_{n \geq 1} \int_0^\infty e^{-t} \frac{t^{n-1}}{(n-1)!} P_t(i, j) dt = \int_0^\infty P_t(i, j) dt.$$

Remarquons ensuite que

$$\int_0^\infty P_t(i, j) dt = \mathbf{E}_i \left(\int_0^\infty \mathbf{1}_{\{X_t=j\}} dt \right) = \sum_{n=0}^\infty \mathbf{E}_i [\mathbf{1}_{\{X_{\tau_n}=j\}} (T_{n+1} - T_n)]$$

car $\mathbf{1}_{\{X_t=j\}} = \sum_{n \geq 0} \mathbf{1}_{\{X_{T_n}=j\}} \mathbf{1}_{\{t \in [T_n, T_{n+1})\}}$, puisque $X_t = \sum_{n \geq 0} X_{T_n} \mathbf{1}_{\{t \in [T_n, T_{n+1})\}}$. Il s'ensuit que

$$\int_0^\infty P_t(i, j) dt = \sum_{n=0}^\infty \mathbf{E}_i (\mathbf{1}_{\{X_{T_n}=j\}} \mathbf{E}_i [T_{n+1} - T_n \mid (X_{T_k})_{k \geq 0}]).$$

Mais conditionnellement à $(X_{T_k})_{k \geq 0}$, on a $T_{n+1} - T_n \sim \text{Exp}(\lambda(X_{T_n}))$. Ainsi,

$$\int_0^\infty P_t(i, j) dt = \sum_{n=0}^\infty \mathbf{E}_i \left[\mathbf{1}_{\{X_{T_n}=j\}} \cdot \frac{1}{\lambda(X_{T_n})} \right] = \frac{1}{\lambda(j)} \sum_{n=0}^\infty Q^n(i, j).$$

Nous avons donc montré (★), ce qui conclut la preuve. $\square$

Remarque. L'expression pour $\Pi^n(i, j)$ en (★★) est immédiate en avec point (ii) de la remarque précédente, et le fait que qu'une somme de n variables aléatoires iid exponentielles de paramètre 1, est de loi Gamma(1, n).

Proposition 5.14. Si $(X_t)_{t \geq 0}$ est un PMS irréductible et récurrent positif, alors il est récurrent.

Preuve. On note π sa probabilité invariante (telle que $\pi P_t = \pi$ pour tout $t \geq 0$), et on doit montrer que $G(i, j) = \infty$, où $G(i, j) = \sum_{n=0}^\infty Q^n(i, j)$. On a, pour tout $j \in E$,

$$\mathbf{E}_\pi \left[\int_0^\infty \mathbf{1}_{\{X_t=j\}} dt \right] = \int_0^\infty (\pi P_t)(j) dt = \infty, \quad \text{i.e.} \quad \sum_{i \in E} \int_0^\infty \pi(i) P_t(i, j) dt = \infty.$$

Mais on a vu dans la preuve précédente que

$$\int_0^\infty P_t(i, j) dt = \frac{1}{\lambda(j)} \sum_{n \geq 0} Q^n(i, j) = \frac{1}{\lambda(j)} G(i, j).$$

Par conséquent, en utilisant le principe du maximum,

$$\int_0^\infty (\pi P_t)(j) dt = \frac{1}{\lambda(j)} \sum_{i \in E} \pi(i) G(i, j) \leq \frac{1}{\lambda(j)} \sum_{i \in E} \pi(i) G(j, j) = \frac{1}{\lambda(j)} G(j, j),$$

puisque π est une mesure de probabilité. Ainsi,

$$G(j, j) \geq \lambda(j) \int_0^\infty (\pi P_t)(j) dt = \infty,$$

et l'état $j \in E$ est récurrent. Comme $j \in E$ était arbitraire, la preuve est terminée. $\square$

Théorème 5.15. Soit $(X_t)_{t \geq 0}$ un PMS irréductible et **récurrent**. Il admet une mesure invariante m , unique à une constante multiplicative près. De plus, on a

1. Pour $i \in E$ fixé, si $S_i = \inf \{t > T_1; X_t = i\}$, $m_i(j) = \mathbf{E}_i [\int_0^{S_i} \mathbf{1}_{\{X_s=j\}} ds]$ est une mesure invariante.
2. Si m est une mesure telle que ν définie par $\nu(j) = \lambda(j)m(j)$ vérifie $\nu Q = \nu$ (i.e. pour tout $i \in E$, $\lambda(i)m(i) = \sum_{j \neq i} \lambda(j)m(j)Q(j, i)$), alors m est invariante.
3. Si $mA = 0$ (i.e. pour tout $i \in E$, $m(i)\lambda(i) = \sum_{j \neq i} m(j)\lambda(j)Q(j, i)$), alors m est invariante.

Remarque. On observe qu'informellement, par Kolmogorov rétrograde, si $mA = 0$, alors pour tout $t \geq 0$, $(mP_t)' = mP_t' = mAP_t = 0$, donc $mP_t = mP_0 = m$. Mais nous avons échangé une somme et une dérivation, donc ce n'est rigoureux que dans le cas où l'espace d'états E est fini.

Preuve. Unicité. Si m est une mesure invariante, $mP_t = m$ pour tout $t \geq 0$. Alors $m\Pi = m$, où $\Pi(i, j) = \int_0^\infty e^{-t} P_t(i, j) dt$ est la transition introduite dans le lemme, car

$$(m\Pi)(j) = \int_0^\infty e^{-t} (mP_t)(j) dt = \int_0^\infty e^{-t} m(j) dt = m(j).$$

Donc m est une mesure invariante de la chaîne de noyau Π . Or d'après les résultats sur les chaînes de Markov, on sait qu'il n'existe qu'une seule telle mesure, à une constante près, car cette chaîne est irréductible et récurrente (par le lemme et car notre PMS est irréductible et récurrent).

Existence et point 1. On fixe $i \in E$ et on montre que la mesure définie au point 1 est invariante. On fixe $t > 0$ et $k \in E$, et on montre que $(m_i P_t)(k) = m_i(k)$. On se rappelle que $(m_i P_t)(k) = \sum_{j \in E} m_i(j) P_t(j, k)$, et on utilise la définition de m_i :

$$(m_i P_t)(k) = \sum_{j \in E} \mathbf{E}_i \left[\int_0^{S_i} \mathbf{1}_{\{X_s=j\}} \mathbf{P}_j(X_t=k) ds \right] = \mathbf{E}_i \left[\int_0^{S_i} \mathbf{P}_{X_s}(X_t=k) ds \right].$$

Mais $\mathbf{P}_{X_s}(X_t=k) = \mathbf{P}_i(X_{t+s}=k \mid \mathcal{F}_s)$ par la propriété de Markov simple. Puis, comme S_i est un temps d'arrêt, $\{S_i > s\} = \{S_i \leq s\}^c \in \mathcal{F}_s$ et

$$\mathbf{1}_{\{S_i > s\}} \mathbf{P}_{X_s}(X_t=k) = \mathbf{1}_{\{S_i > s\}} \mathbf{P}_i(X_{t+s}=k \mid \mathcal{F}_s) = \mathbf{E}_i(\mathbf{1}_{\{S_i > s, X_{t+s}=k\}} \mid \mathcal{F}_s).$$

Donc

$$\begin{aligned} (m_i P_t)(k) &= \mathbf{E}_i \left[\int_0^\infty \mathbf{1}_{\{S_i > s\}} \mathbf{P}_{X_s}(X_t=k) ds \right] \\ &= \mathbf{E}_i \left[\int_0^\infty \mathbf{E}_i(\mathbf{1}_{\{S_i > s, X_{t+s}=k\}} \mid \mathcal{F}_s) ds \right] \\ &= \mathbf{E}_i \left[\int_0^\infty \mathbf{1}_{\{S_i > s, X_{t+s}=k\}} ds \right] \\ &= \mathbf{E}_i \left[\int_0^{S_i} \mathbf{1}_{\{X_{t+s}=k\}} ds \right]. \end{aligned}$$

En changeant de variable, $s \rightarrow u = t + s$,

$$\begin{aligned} (m_i P_t)(k) &= \mathbf{E}_i \left[\int_t^{t+S_i} \mathbf{1}_{\{X_u=k\}} du \right] \\ &= \underbrace{\mathbf{E}_i \left[\int_t^{S_i} \mathbf{1}_{\{X_u=k\}} du \right]}_A + \underbrace{\mathbf{E}_i \left[\int_{S_i}^{S_i+t} \mathbf{1}_{\{X_u=k\}} du \right]}_B. \end{aligned}$$

En utilisant la propriété de Markov forte en S_i avec $\tilde{X}_t = X_{S_i+t}$,

$$\begin{aligned}
 B &= \mathbf{E}_i \left[\int_{S_i}^{S_i+t} \mathbf{E}_i [\mathbf{1}_{\{X_u=k\}} \mid \mathcal{F}_{S_i}] du \right] \\
 &= \mathbf{E}_i \left[\int_0^t \mathbf{E}_i [\mathbf{1}_{\{X_{S_i+u}=k\}} \mid \mathcal{F}_{S_i}] du \right] \\
 &= \mathbf{E}_i \left[\int_0^t \mathbf{E}_i [\mathbf{1}_{\{\tilde{X}_u=k\}} \mid \mathcal{F}_{S_i}] du \right] \\
 &= \mathbf{E}_i \left[\int_0^t \mathbf{E}_{X_{S_i}} [\mathbf{1}_{\{X_u=k\}}] du \right] \\
 &= \mathbf{E}_i \left[\int_0^t \mathbf{E}_i [\mathbf{1}_{\{X_u=k\}}] du \right] \\
 &= \mathbf{E}_i \left[\int_0^t \mathbf{1}_{\{X_u=k\}} du \right].
 \end{aligned}$$

Enfin, on voit que

$$(m_i P_t)(k) = A + B = \mathbf{E}_i \left[\int_t^{S_i} \mathbf{1}_{\{X_u=k\}} du \right] + \mathbf{E}_i \left[\int_0^t \mathbf{1}_{\{X_u=k\}} du \right] = \mathbf{E}_i \left[\int_0^{S_i} \mathbf{1}_{\{X_u=k\}} du \right] = m_i(k).$$

Point 2. On fixe $i \in E$, on considère la mesure invariante m_i construite au point 1, on pose $\nu(j) = \lambda(j)m_i(j)$ et on montre que $\nu Q = \nu$ et on conclura ainsi : soit $\tilde{m}$ une mesure telle que, pour $\tilde{\nu}(j) = \lambda(j)\tilde{m}(j)$, on ait $\tilde{\nu}Q = \tilde{\nu}$, alors, par unicité de la mesure invariante de Q , on aura que $\tilde{\nu}$ est proportionnelle à ν , et donc $\tilde{m}$ est proportionnelle à m_i , et donc que $\tilde{m}$ est invariante.

Avec $\tau_i = \inf\{n \geq 1 : X_{T_n} = i\}$, on a $S_i = T_{\tau_i}$. Donc, en se rappelant que X_t est constant sur les intervalles $[T_n, T_{n+1}[$, ce qui s'écrit aussi $X_t = \sum_{n \geq 0} X_{T_n} \mathbf{1}_{\{t \in [T_n, T_{n+1}[\}}$, on obtient

$$X_t \mathbf{1}_{\{t < S_i\}} = \sum_{n=0}^{\tau_i-1} X_{T_n} \mathbf{1}_{\{t \in [T_n, T_{n+1}[\}}.$$

Puis

$$m_i(j) = \mathbf{E}_i \left[\int_0^{S_i} \mathbf{1}_{\{X_t=j\}} dt \right] = \mathbf{E}_i \left[\sum_{n=0}^{\tau_i-1} (T_{n+1} - T_n) \mathbf{1}_{\{X_{T_n}=j\}} \right] = \frac{1}{\lambda(j)} \mathbf{E}_i \left[\sum_{n=0}^{\tau_i-1} \mathbf{1}_{\{X_{T_n}=j\}} \right],$$

où pour la seconde égalité nous avons utilisé que sachant $(X_{T_k})_{k \geq 0}$, $(T_{n+1} - T_n) \sim \text{Exp}(\lambda(X_{T_n}))$, et que τ_i est bien sûr $\sigma((X_{T_k})_{k \geq 0})$ -mesurable. Enfin,

$$\nu(j) = \lambda(j)m_i(j) = \mathbf{E}_i \left[\sum_{n=0}^{\tau_i-1} \mathbf{1}_{\{X_{T_n}=j\}} \right],$$

qui est invariante pour la chaîne $(X_{T_n})_{n \geq 0}$ (cf Lemme 4.12) et donc vérifie $\nu = \nu Q$.

Point 3. C'est juste une réécriture du point 2 en utilisant le générateur. □

Proposition 5.16. Si X_t est un PMS récurrent positif, de probabilité invariante π , alors pour toute $f : E \rightarrow \mathbb{R}$ (bornée ou positive), pour tout $i \in E$, $\mathbf{P}_i$ -p.s.,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(X_s) ds = \int_E f d\pi.$$

Preuve. Soit $i \in E$ fixé et

$$S_i = \inf\{t > T_1 : X_t = i\} = \inf\{t > 0 : X_t = i \text{ et } \exists s \in [0, t] \text{ t.q. } X_s \neq i\}.$$

Définissons, par récurrence, $\tau_0 = 0$, $\tau_1 = S_i$, puis

$$\tau_{n+1} = \inf\{t > \tau_n : X_t = i \text{ et } \exists s \in [\tau_n, t] \text{ t.q. } X_s \neq i\}.$$

En posant $\tilde{X}_t = X_{\tau_n+t}$, on a

$$\int_0^{\tilde{S}_i} f(\tilde{X}_s) ds = \int_{\tau_n}^{\tau_{n+1}} f(X_s) ds,$$

pour toute fonction $f : E \rightarrow \mathbb{R}_+$. On en déduit, par la propriété de Markov forte en τ_n , que $\tilde{X}$ a la même loi que X et est indépendant de $\mathcal{F}_{\tau_n}$. Une récurrence simple montre alors que les v.a. $(Z_n)_{n \geq 0}$ données par

$$Z_n = \int_{\tau_n}^{\tau_{n+1}} f(X_s) ds,$$

sont i.i.d., et que par conséquent, la suite $(\tau_n)_{n \geq 0}$ fait de $(X_t)_{t \geq 0}$ un processus régénératif.

Il résulte donc du théorème ergodique pour les processus régénératifs (théorème 3.2) que $\mathbf{P}_i$ -p.s.,

$$\lim_{t \rightarrow \infty} t^{-1} \int_0^t f(X_s) ds = \int_E f d\pi, \quad \text{où} \quad \pi(j) = \frac{1}{\mathbf{E}_i[S_i]} \mathbf{E}_i \left[\int_0^{S_i} \mathbf{1}_{\{X_s=j\}} ds \right].$$

Par le théorème précédent, π est bien l'unique probabilité invariante (elle est invariante, et bien sûr, $\sum_{j \in E} \pi(j) = \mathbf{E}_i[S_i] / \mathbf{E}_i[S_i] = 1$). □

Résumé important. (i) Si Q est irréductible et récurrente, alors le PMS associé à Q et $(\lambda(i))_{i \in E}$ admet une unique mesure invariante (à constante près) m , caractérisée par $mA = 0$.

(ii) Il est possible, si Q n'est **pas récurrente**, de trouver une probabilité m telle que $mA = 0$ sans que le PMS (ni Q) soit récurrent (une telle probabilité m n'est alors pas invariante!).

(iii) Si une probabilité m est telle que $mA = 0$ et $\sum_{i \in E} m(i) \lambda(i) < \infty$, alors m est une probabilité invariante du PMS, qui est donc récurrent positif. En effet, on a alors que $\nu(i) = \lambda(i) m(i)$ est invariante pour Q . Comme ν est une mesure finie, on peut trouver une probabilité invariante pour Q , la chaîne associée est récurrente positive et par conséquent récurrente. Donc le PMS est récurrent et on peut appliquer le théorème 5.15: m est invariante.

(iv) Il est possible que Q soit récurrente nulle et le PMS récurrent positif (Exemple 5.1).

(v) Il est possible que Q soit récurrente positive et le PMS est récurrent nul (Exemple 5.2).

Exemple 5.1 (Un PMS récurrent positif avec une chaîne récurrente nulle). On considère $E = \mathbb{Z}$, $Q(i, i-1) = Q(i, i+1) = 1/2$. On a vu (c'est la marche aléatoire simple symétrique) que la chaîne de transition Q est récurrente nulle, de mesure invariante $\nu(i) = 1$ pour tout $i \in \mathbb{Z}$. Si $\lambda(i) = (1 + i^2)$, alors $m(i) = \nu(i) / \lambda(i) = 1 / (1 + i^2)$ vérifie $mA = 0$ (pour le PMS de caractéristiques Q et λ), c'est donc une mesure invariante finie du PMS, et $\pi(i) = a / (1 + i^2)$ (où $a = 1 / \sum_{i \in \mathbb{Z}} (1 + i^2)^{-1}$) est une probabilité invariante du PMS, qui est donc récurrent positif.

Exemple 5.2 (Un PMS récurrent nul avec une chaîne récurrente positive). On considère n'importe quelle matrice de transition irréductible et récurrente positive Q , sur un ensemble dénombrable infini E , de probabilité invariante ν . Alors le PMS de caractéristiques Q et $\lambda(i) = \nu(i)$ est récurrent (puisque Q l'est), et sa mesure invariante est donnée par $m(i) = \nu(i)/\lambda(i) = 1$, qui n'est pas sommable.

5.6 Processus de naissance et mort

Définition (Processus de naissance et mort). On appelle **processus de naissance et mort** (PNM) tout PMS à valeurs dans $\mathbb{N}$ tel que $Q(i, j) = 0$ si $j \notin \{i-1, i+1\}$. On a donc aussi $A(i, j) = 0$ si $j \notin \{i-1, i, i+1\}$.

En définissant $\alpha_i = A(i, i-1)$ pour tout $i \geq 1$, $\beta_i = A(i, i+1)$ pour tout $i \geq 0$. On a alors

- (i) $\lambda(0) = \beta_0$ et $Q(0, 1) = 1$;
- (ii) pour $i \geq 1$, $\lambda(i) = \alpha_i + \beta_i$, $Q(i, i-1) = \frac{\alpha_i}{\alpha_i + \beta_i}$, $Q(i, i+1) = \frac{\beta_i}{\alpha_i + \beta_i}$.

L'idée est que le PMS $(X_t)_{t \geq 0}$ modélise une population : X_t est le nombre d'individus vivants à l'instant t . Pour tout $i \geq 1$, α_i est le *taux de mort* quand i individus sont vivants, et β_i est le *taux de naissance* quand i individus sont vivants. Donc (c'est lié au lemme des deux réveils) quand on est dans l'état i , on saute avec taux $\lambda(i) = \beta_i + \alpha_i$, et c'est une naissance (donc on saute de 1) avec probabilité $\frac{\beta_i}{\alpha_i + \beta_i}$ ou c'est une mort (donc on saute de -1) avec probabilité $\frac{\alpha_i}{\alpha_i + \beta_i}$.

Remarque. Il découle du cas à temps discret qu'un PNM est irréductible si et seulement si $\alpha_i > 0$ pour tout $i \geq 1$ et $\beta_i > 0$ pour tout $i \geq 0$.

Théorème 5.17. Un PNM irréductible est

- (i) récurrent si et seulement si

$$\sum_{i \geq 1} \frac{\alpha_1 \alpha_2 \cdots \alpha_i}{\beta_1 \beta_2 \cdots \beta_i} = \infty.$$

- (ii) récurrent positif si et seulement si

$$\sum_{i \geq 1} \frac{\alpha_1 \alpha_2 \cdots \alpha_i}{\beta_1 \beta_2 \cdots \beta_i} = \infty \quad \text{et} \quad S = 1 + \sum_{i \geq 1} \frac{\beta_0 \beta_1 \cdots \beta_{i-1}}{\alpha_1 \alpha_2 \cdots \alpha_i} < \infty.$$

Dans le dernier cas, sa probabilité invariante est alors

$$\pi(0) = \frac{1}{S} \quad \text{et} \quad \pi(i) = \frac{\beta_0 \beta_1 \cdots \beta_{i-1}}{S \alpha_1 \alpha_2 \cdots \alpha_i} \quad \text{pour } i \geq 1.$$

Preuve. (i) Par définition, le PMS est récurrent ssi la chaîne de transition Q l'est. Mais cette chaîne est un PNM à temps discret. On peut donc appliquer le théorème 4.23, et on voit que notre PNM est récurrent ssi

$$\sum_{i \geq 1} \frac{q_1 q_2 \cdots q_i}{p_1 p_2 \cdots p_i} = \infty,$$

où $q_i = Q(i, i-1)$ et $p_i = Q(i, i+1)$. La conclusion s'ensuit immédiatement, puisque $q_i/p_i = \alpha_i/\beta_i$.

(ii) Supposons $S < \infty$. Le processus étant récurrent, il suffit de vérifier que π , qui est bien sûr une probabilité, vérifie est invariante, c'est-à-dire que $\pi(i)\lambda(i) = \sum_{j \neq i} \pi(j)\lambda(j)Q(j, i)$ pour tout $i \in \mathbb{N}$.

Si $i = 0$, on veut

$$\pi(0)\beta_0 = \pi(1)(\alpha_1 + \beta_1) \frac{\alpha_1}{\alpha_1 + \beta_1},$$

i.e. $\pi(0)\beta_0 = \pi(1)\alpha_1$, ce qui est vrai.

Si $i \geq 1$, il faut

$$\pi(i)(\alpha_i + \beta_i) = \pi(i-1)(\alpha_{i-1} + \beta_{i-1}) \frac{\beta_{i-1}}{\alpha_{i-1} + \beta_{i-1}} + \pi(i+1)(\alpha_{i+1} + \beta_{i+1}) \frac{\alpha_{i+1}}{\alpha_{i+1} + \beta_{i+1}},$$

soit encore $\pi(i)(\alpha_i + \beta_i) = \pi(i-1)\beta_{i-1} + \pi(i+1)\alpha_{i+1}$, c'est vrai aussi. Par conséquent π est une probabilité invariante et le processus est récurrent positif.

Si au contraire $S = \infty$, on vérifie que la mesure m donnée par $m(0) = 1$ et $m(i) = \frac{\beta_0 \beta_1 \dots \beta_{i-1}}{\alpha_1 \alpha_2 \dots \alpha_i}$ pour $i \geq 1$ est invariante; c'est exactement la même preuve. Comme m n'est pas sommable, le PMS n'est pas récurrent positif (). $\square$

Exemple 5.3 (Une probabilité π non-invariante telle que $\pi A = 0$). Prenons $\alpha_n = 3^n, \beta_n = 2 \cdot 3^n$. Alors $\sum_{n=1}^{\infty} \frac{\alpha_1 \dots \alpha_n}{\beta_1 \dots \beta_n} = \sum_{n=1}^{\infty} \frac{1}{2^n} < +\infty$, donc le processus n'est pas récurrent. Pourtant, $S < \infty$ et la probabilité π du point (ii) vérifie bien $\pi A = 0$. Elle n'est pas invariante, puisque sinon, le processus serait récurrent positif. Ce PNM explose p.s., car la proposition suivante s'applique.

Proposition 5.18. On considère un PNM $(X_t)_{t \geq 0}$ irréductible. Si

$$R = \sum_{n=0}^{\infty} \left(\frac{1}{\beta_n} + \frac{\alpha_n}{\beta_n \beta_{n-1}} + \frac{\alpha_n \alpha_{n-1}}{\beta_n \beta_{n-1} \beta_{n-2}} + \dots + \frac{\alpha_n \dots \alpha_1}{\beta_n \dots \beta_0} \right) < \infty,$$

alors il y a p.s. explosion, c'est-à-dire que $\mathbf{P}_0(\lim_n T_n < \infty) = 1$.

Remarque. On montre (cf appendice) que si $R = \infty$, il n'y a jamais d'explosion: $\mathbf{P}_i(\lim_n T_n < \infty) = 0$ pour tout $i \in \mathbb{N}$.

Preuve. On fixe $a \in \mathbb{N}$ grand, et on pose $u_a(i) = \mathbf{E}_i[S_a]$, où $S_a = \inf\{t \geq 0 : X_t = a\}$. On a $u_a(a) = 0$ et, pour $i \in \{1, \dots, a-1\}$, en utilisant la propriété de Markov forte en T_1 avec $\tilde{X}_t = X_{T_1+t}$, donc $S_a = T_1 + \tilde{S}_a$, on a

$$\begin{aligned} u_a(i) &= \mathbf{E}_i[T_1 + \mathbf{E}_i[\tilde{S}_a | \mathcal{F}_{T_1}]] \\ &= \mathbf{E}_i[T_1] + \mathbf{E}_1[\mathbf{E}_{X_{T_1}}[S_a]] \\ &= \frac{1}{\alpha_i + \beta_i} + \frac{\alpha_i}{\alpha_i + \beta_i} u_a(i-1) + \frac{\beta_i}{\alpha_i + \beta_i} u_a(i+1), \end{aligned}$$

soit encore $(\alpha_i + \beta_i)u_a(i) = 1 + \alpha_i u_a(i-1) + \beta_i u_a(i+1)$, que l'on peut écrire de manière à faire apparaître une récurrence:

$$u_a(i) - u_a(i+1) = \frac{1}{\beta_i} + \frac{\alpha_i}{\beta_i} [u_a(i-1) - u_a(i)], \quad i \in \{1, \dots, a-1\}.$$

De même, on a $u_a(0) = \mathbf{E}_0[T_1] + \mathbf{E}_1[S_a] = 1/\beta_0 + u_a(1)$.

En travaillant un peu, on r sout la r currence de proche en proche

$$\begin{aligned}
u_a(0) - u_a(1) &= \frac{1}{\beta_0}, \\
u_a(1) - u_a(2) &= \frac{1}{\beta_1} + \frac{\alpha_1}{\beta_1\beta_0}, \\
u_a(2) - u_a(3) &= \frac{1}{\beta_2} + \frac{\alpha_2}{\beta_2\beta_1} + \frac{\alpha_2\alpha_1}{\beta_2\beta_1\beta_0}, \\
&\vdots \\
u_a(a-1) - u_a(a) &= \frac{1}{\beta_{a-1}} + \frac{\alpha_{a-1}}{\beta_{a-1}\beta_{a-2}} + \dots + \frac{\alpha_{a-1}\dots\alpha_1}{\beta_{a-1}\dots\beta_0}.
\end{aligned}$$

En sommant tout cela, comme $u_a(a) = 0$, on trouve

$$\mathbf{E}_0[S_a] = u_a(0) = \sum_{n=0}^{a-1} \left(\frac{1}{\beta_n} + \frac{\alpha_n}{\beta_n\beta_{n-1}} + \frac{\alpha_n\alpha_{n-1}}{\beta_n\beta_{n-1}\beta_{n-2}} + \dots + \frac{\alpha_n\dots\alpha_1}{\beta_n\dots\beta_0} \right).$$

Comme par hypoth se la somme du membre de droite est finie, on en d duit que $\lim_{a \rightarrow \infty} \mathbf{E}_0[S_a] < \infty$. Mais bien s r, il faut au moins n sauts pour aller en n en partant de z ro et on a $T_n \leq S_n$ pour tout $n \geq 1$ sous $\mathbf{P}_0$. Ainsi, $\lim_{n \rightarrow \infty} \mathbf{E}_0[T_n] < \infty$. Comme $n \mapsto T_n$ est croissant, on conclut par le lemme de Fatou que $\mathbf{E}_0[\lim_{n \rightarrow \infty} T_n] \leq \lim_n \mathbf{E}_0[T_n] < \infty$, ce qui implique que $\mathbf{P}_0(\lim_n T_n < \infty) = 1$. $\square$

Files d'attente et autres applications

Ce chapitre contient un certain nombre d'applications à ces processus concrets qui interviennent souvent en modélisation, notamment des (réseaux de) files d'attente, et des modèles de dynamique des population.

6.1 Généralités sur les files d'attente

Nous allons étudier quelques modèles de files d'attentes plus complexe que le modèle très simple de la section 2.4. L'objectif principal est de comprendre comment trouver le générateur, i.e. les taux d'événements $(\lambda(i))_{i \in E}$ et la matrice $(Q(i, j))_{i, j \in E}$ dans des situations concrètes.

On modélise une file d'attente de la manière suivante. Une file d'attente est constituée de

- clients qui arrivent de l'extérieur pour rejoindre cette file,
- de guichets où les clients vont se faire servir par des serveurs;
- dans certains cas les clients attendent dans une salle d'attente de capacité limitée;
- un client servi disparaît du système;
- les instants d'arrivée des clients et les temps de service sont aléatoires.

Par conséquent, une file d'attente est décrite par

- la loi d'interarrivée des clients,
- la loi des temps de service,
- le nombre de serveurs, et
- la longueur maximale de la file (égale à la taille de la salle d'attente éventuelle).

Nous supposons toujours ici que les interarrivées sont des variables aléatoires indépendantes et de même loi, indépendantes des temps de service, eux mêmes indépendants et de même loi. Pour les files simples, on utilise les notations introduites par Kendall:

Loi d'interarrivée / Loi de service / Nombre de serveurs (/ Longueur maximale).

Les lois sont notées symboliquement: M lorsqu'elles sont exponentielles (M pour Markov), G (G pour Général) sinon. On ne spécifie pas la longueur maximale de la file lorsqu'elle est infinie. Nous nous limiterons ici à l'étude de files markoviennes.

Exemple 6.1. Par exemple une file $M/M/s$ est une file d'attente à s guichets, telle que le flot d'arrivée des clients est poissonnien et les temps de service exponentiels, sans restriction sur la taille de la file d'attente.

La question essentielle est de savoir si la taille de la file va exploser ou au contraire s'équilibrer. Dans ce dernier cas, il peut être intéressant de calculer la taille moyenne de la file, la loi du temps d'attente d'un client, etc...

6.2 La file $M/M/s$.

Il y a $s \in \mathbb{N} \cup \{\infty\}$ serveurs. Les interarrivées sont des v.a. exponentielles de paramètre α et les temps de service des v.a. exponentielles de paramètre μ , toutes indépendantes. On note

X_t = nombre de clients dans file+caisse à l'instant t .

et on considère $\mathcal{F}_t = \sigma(X_s, s \leq t)$.

Théorème 6.1. Le processus $(X_t)_{t \geq 0}$ est un PMS à valeurs dans $\mathbb{N}$. De plus, on a pour tout $i \geq 0$, $\lambda(i) = \alpha + \min\{i, s\}\mu$ et

$$Q(i, i-1) = \frac{\min\{i, s\}\mu}{\alpha + \min\{i, s\}\mu} \quad \text{et} \quad Q(i, i+1) = \frac{\alpha}{\alpha + \min\{i, s\}\mu}.$$

Noter que $Q(0, -1)$ (qui n'existe pas vraiment puisque $E = \mathbb{N}$) vaut bien 0. Dans tous les autres cas, $Q(i, j) = 0$.

Preuve. Etape 1. Le processus $(X_t)_{t \geq 0}$ est un PMS car les lois en jeu sont exponentielles (sans mémoire). Si on connaît $(X_s)_{s \in [0, t]}$, alors ce qui se passe après t ne dépend que de X_t .

En effet, si $X_t = i \geq 0$, il y a précisément

- $k = \min\{s, i\}$ clients en train de se faire servir, et
- $i - k$ (qui est nul si $i \leq s$) en train d'attendre.

Pour chaque client au service, on sait depuis quand il est arrivé, quand il a commencé à se faire servir, mais cela n'apporte aucune information : comme la loi de service $\text{Exp}(\mu)$ est sans mémoire, son temps de service résiduel est encore de loi $\text{Exp}(\mu)$. Un client est en train d'arriver, et on sait depuis combien de temps on l'attend, mais, la loi des interarrivées $\text{Exp}(\alpha)$ étant sans mémoire, notre temps résiduel d'attente sera encore de loi $\text{Exp}(\alpha)$. Le futur après t ne dépend donc de $(X_s)_{s \in [0, t]}$ que par le fait que $X_t = i$.

Etape 2. On détermine le générateur. On pose $T_1 = \inf\{t \geq 0 : X_t \neq X_0\}$, et on cherche $Q(i, j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i[T_1]$ pour $i, j \in \mathbb{N}$. On suppose donc que $X_0 = i$.

Si $i = 0$, soit T le temps d'arrivée du prochain client. Par hypothèse, $T \sim \text{Exp}(\alpha)$. On a forcément $T_1 = T$ et $X_{T_1} = 1$. Du coup, on a bien $\lambda(0) = \alpha = \alpha + \min\{0, s\}\mu$ et $Q(0, 1) = 1$.

Si $i \geq 1$, il y a $k = \min\{s, i\}$ clients au service, notons $S_1, \dots, S_k$ leurs temps de service, qui sont i.i.d. $\text{Exp}(\mu)$. Il y a un client en train d'arriver, notons $T \sim \mathcal{E}(\alpha)$ son temps d'arrivée. Notons $S = \min\{S_1, \dots, S_k\}$, dont la loi est $\text{Exp}(k\mu)$ (par le lemme des réveils). Alors $T_1 = \min\{S, T\}$ et on a $X_{T_1} = i + 1$ si $T < S$ et $X_{T_1} = i - 1$ si $S < T$. Par le lemme des réveils, on voit que $T_1 \sim \text{Exp}(\alpha + k\mu)$, donc $\lambda(i) = 1/\mathbf{E}_i[T_1] = \alpha + k\mu$

et que $Q(i, i+1) = \mathbf{P}_i(X_{T_1} = i+1) = \mathbf{P}(T < S) = \alpha/(\alpha + k\mu)$ et que $Q(i, i-1) = \mathbf{P}_i(X_{T_1} = i-1) = \mathbf{P}(S < T) = k\mu/(\alpha + k\mu)$. $\square$

En utilisant les résultats généraux sur les processus de naissance et mort (PNM) et le fait qu'on a ici (avec les notations de la section 5.6)

- $\alpha_i = \lambda(i)Q(i, i-1) = \min\{i, s\}\mu$ pour $i \geq 1$ et,
- $\beta_i = \lambda(i)Q(i, i+1) = \alpha$ pour $i \geq 0$,

on peut montrer la proposition suivante.

Proposition 6.2. Le PMS $(X_t)_{t \geq 0}$ est récurrent ssi $\alpha \leq s\mu$. Il est récurrent positif ssi $\alpha < s\mu$ et sa probabilité invariante π est (moche mais) calculable explicitement. Toutefois, on a

- $\pi(i) = (1 - \alpha/\mu)(\alpha/\mu)^i$ pour tout $i \in \mathbb{N}$ si $s = 1$ (et $\alpha < \mu$),
- $\pi(i) = e^{-\alpha/\mu}(\alpha/\mu)^i / i!$ pour tout $i \in \mathbb{N}$ si $s = \infty$.

Le processus est récurrent si la file n'explose pas (en temps infini) : il faut pour cela que le taux d'arrivée α soit inférieur au taux de départ maximal μs (taux de départ par serveur multiplié par le nombre de serveurs).

Kolmogorov forward. Vérifier que KF s'écrit, pour ce processus, pour tout $i \geq 0$,

$$\begin{cases} P'_t(i, 0) = -\alpha P_t(i, 0) + \mu P_t(i, 1), \\ P'_t(i, j) = -(\alpha + \mu \min\{j, s\})P_t(i, j) + \alpha P_t(i, j-1) + \mu \min\{j+1, s\}P_t(i, j+1), \end{cases}$$

la seconde ligne étant valide pour tout $j \geq 1$. C'est "limpide" : par exemple si $j \geq 1$,

- le terme $-(\alpha + \mu \min\{j, s\})P_t(i, j)$ exprime qu'on part de j avec taux $\alpha + \mu \min\{j, s\}$;
- le terme $+\alpha P_t(i, j-1)$ exprime qu'on passe de $j-1$ à j à taux α ;
- le terme $+\mu \min\{j+1, s\}P_t(i, j+1)$ exprime qu'on passe de $j+1$ à j à taux $\mu \min\{j+1, s\}$.

On peut s'amuser à sommer (sans trop se poser de questions) les égalités ci-dessus. Par exemple, en utilisant que $\mathbf{E}_i[X_t] = \sum_{j \geq 1} j P_t(i, j)$, on peut écrire

$$\begin{aligned} \frac{d}{dt} \mathbf{E}_i[X_t] &= \sum_{j \geq 1} j P'_t(i, j) \\ &= \sum_{j \geq 1} (\alpha + \mu \min\{j, s\}) j P_t(i, j) + \sum_{j \geq 1} \alpha j P_t(i, j-1) + \sum_{j \geq 1} \mu \min\{j+1, s\} j P_t(i, j+1) \\ &= -\mathbf{E}_i[(\alpha + \mu \min\{X_t, s\})X_t] + \sum_{\ell \geq 0} \alpha(\ell+1)P_t(i, \ell) + \sum_{\ell \geq 2} \mu \min\{\ell, s\}(\ell-1)P_t(i, \ell) \\ &= -\mathbf{E}_i[(\alpha + \mu \min\{X_t, s\})X_t] + \mathbf{E}_i[\alpha(X_t+1)] + \mathbf{E}_i[\mu \min\{X_t, s\}(X_t-1)] \\ &= \alpha - \mu \mathbf{E}_i[\min\{X_t, s\}]. \end{aligned}$$

Voir l'appendice pour une preuve rigoureuse. Quand $s = \infty$, cela donne

$$\frac{d}{dt} \mathbf{E}_i[X_t] = \alpha - \mu \mathbf{E}_i[X_t],$$

qu'on peut résoudre explicitement. Comme $\mathbf{E}_i[X_0] = i$, on trouve

$$\mathbf{E}_i[X_t] = \frac{\alpha}{\mu} + \left(i - \frac{\alpha}{\mu}\right)e^{-\mu t}.$$

On pourra comparer $\lim_{t \rightarrow \infty} \mathbf{E}_i[X_t]$ avec la moyenne de la probabilité invariante du processus, qui est Poisson(α/μ) dans ce cas.

6.3 La file M/M/1/k.

Il y a un unique serveur. Les interarrivées sont des v.a. exponentielles de paramètre α et les temps de service des v.a. exponentielles de paramètre μ , toutes indépendantes. Par contre, la salle d'attente est limitée : si un client qui arrive trouve k personnes dans la salle d'attente, il s'en va immédiatement. On note

X_t = nombre de clients dans file+caisse à l'instant t .

Théorème 6.3. Le processus $(X_t)_{t \geq 0}$ est un PMS à valeurs dans $\{0, \dots, k+1\}$, et on a

$$\lambda(i) = \begin{cases} \alpha & \text{si } i = 0 \\ \alpha + \mu & \text{si } i \in \{1, \dots, k\} \\ \mu & \text{si } i = k+1 \end{cases} \quad \text{et} \quad Q(i, j) = \begin{cases} 1 & \text{si } i = 0, j = 1 \\ \frac{\alpha}{\alpha + \mu} & \text{si } i \in \{1, \dots, k\}, j = i+1 \\ \frac{\mu}{\alpha + \mu} & \text{si } i \in \{1, \dots, k\}, j = i-1 \\ 1 & \text{si } i = k+1, j = k. \end{cases}$$

Dans tous les autres cas, $Q(i, j) = 0$.

Preuve. Etape 1. C'est un PMS car les lois en jeu sont exponentielles (sans mémoire). Si on connaît $(X_s)_{s \in [0, t]}$, ce qui se passe après t ne dépend que de X_t . En effet, supposons que $X_t = i$.

Si $i = 0$, un client est en train d'arriver, et on sait depuis combien de temps on l'attend, mais, la loi des interarrivées $\text{Exp}(\alpha)$ étant sans mémoire, notre temps résiduel d'attente sera encore de loi $\text{Exp}(\alpha)$. Le futur après t ne dépend donc de $(X_s)_{s \in [0, t]}$ que par le fait que $X_t = 0$.

Si $i \in \{1, \dots, k\}$, un client est en train d'arriver, et on sait depuis combien de temps on l'attend, mais, la loi des interarrivées $\text{Exp}(\alpha)$ étant sans mémoire, notre temps résiduel d'attente sera encore de loi $\text{Exp}(\alpha)$. Il y a précisément 1 client en train de se faire servir, on sait depuis quand il est arrivé, quand il a commencé à se faire servir, mais cela n'apporte aucune information : comme la loi de service $\text{Exp}(\mu)$ est sans mémoire, son temps de service résiduel est encore de loi $\text{Exp}(\mu)$. Le futur après t ne dépend donc de $(X_s)_{s \in [0, t]}$ que par le fait que $X_t = i$.

Si $i = k+1$, aucun client ne peut arriver. Il y a précisément 1 client en train de se faire servir, on sait depuis quand il est arrivé, quand il a commencé à se faire servir, mais cela n'apporte aucune information : comme la loi de service $\text{Exp}(\mu)$ est sans mémoire, son temps de service résiduel est encore de loi $\text{Exp}(\mu)$. Le futur après t ne dépend donc de $(X_s)_{s \in [0, t]}$ que par le fait que $X_t = k+1$.

Etape 2. On détermine le générateur. On pose donc $T_1 = \inf\{t \geq 0 : X_t \neq X_0\}$, et on cherche $Q(i, j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i[T_1]$. On suppose donc que $X_0 = i$.

Si $i = 0$, soit T le temps d'arrivée du prochain client. Par hypothèse, $T \sim \text{Exp}(\alpha)$. On a forcément $T_1 = T$ et $X_{T_1} = 1$. Du coup, on a bien $Q(0, 1) = 1$ et $\lambda(0) = \alpha$.

Si $i = k+1$. Aucun client ne peut arriver (sans être rejeté). Il y a un client au service, soit S son temps de service, qui suit la loi $\mathcal{E}(\mu)$. On a bien sûr $S \sim \text{Exp}(\mu)$. Forcément, $T_1 = S$ et $X_{T_1} = k$. Du coup, on a bien $Q(k+1, k) = 1$ et $\lambda(k+1) = \mu$.

Si $i \in \{1, \dots, k\}$, alors on a $i-1$ clients dans la file et 1 au service, soit S son temps de service, on a $S \sim \text{Exp}(\mu)$. Soit aussi T le temps d'arrivée du prochain client, on a $T \sim \text{Exp}(\alpha)$. Ainsi, $T_1 = \min\{S, T\}$, et $X_{T_1} = i+1$ si $T < S$ et $X_{T_1} = i-1$ si $S < T$. Par le lemme des réveils, on voit que $T_1 \sim \text{Exp}(\alpha + \mu)$, donc

$\lambda(i) = 1/\mathbf{E}_i[T_1] = \alpha + \mu$ et que $Q(i, i+1) = \mathbf{P}_i(X_{T_1} = i+1) = \mathbf{P}(T < S) = \alpha/(\alpha + \mu)$ et que $Q(i, i-1) = \mathbf{P}_i(X_{T_1} = i-1) = \mathbf{P}(S < T) = \mu/(\alpha + \mu)$. $\square$

6.4 Interlude sur l'équation de Kolmogorov forward

Soit $(X_t)_{t \geq 0}$ un PMS à valeurs dans E , de caractéristiques $(\lambda(i))_{i \in E}$ et $(Q(i, j))_{i, j \in E}$. Soit $\phi : E \rightarrow \mathbb{R}$. On a *a priori* la formule suivante :

$$\frac{d}{dt} \mathbf{E}_i[\phi(X_t)] = \sum_{j \in E} \mathbf{E}_i[\lambda(X_t)(\phi(j) - \phi(X_t))Q(X_t, j)]. \quad (6.1)$$

Attention, cette formule n'est pas toujours vraie : on intervertit sans justification une dérivée et une somme. Le résultat n'est en fait pas toujours vrai. Voir l'appendice pour une justification rigoureuse dans un cas particulier. Tout de même, tout est correct si ϕ est à support fini (i.e. $\phi = 0$ en dehors d'une partie finie de E).

On rappelle l'équation de Kolmogorov forward :

$$\frac{d}{dt} P_t(i, j) = -P_t(i, j)\lambda(j) + \sum_{k \in E} P_t(i, k)\lambda(k)Q(k, j).$$

Donc

$$\begin{aligned} \frac{d}{dt} \mathbf{E}_i[\phi(X_t)] &= \sum_{j \in E} \phi(j) P'_t(i, j) \\ &= - \sum_{j \in E} P_t(i, j)\lambda(j)\phi(j) + \sum_{j \in E} \sum_{k \in E} P_t(i, k)\lambda(k)Q(k, j)\phi(k) \\ &= -\mathbf{E}_i[\lambda(X_t)\phi(X_t)] + \sum_{j \in E} \mathbf{E}_i[\lambda(X_t)Q(X_t, j)\phi(X_t)] \\ &= - \sum_{j \in E} \mathbf{E}_i[\lambda(X_t)\phi(X_t)Q(X_t, j)] + \sum_{j \in E} \mathbf{E}_i[\lambda(X_t)Q(X_t, j)\phi(X_t)]. \end{aligned}$$

6.5 Un réseau de files d'attentes

Des clients arrivent dans un magasin suivant un processus de Poisson de paramètre $\alpha > 0$ (autrement dit, les durées entre les arrivées sont i.i.d. $\text{Exp}(\alpha)$). Ce magasin dispose de s_1 vendeurs, s_2 caissiers et s_3 manutentionnaires, avec $s_1, s_2, s_3 \in \mathbb{N} \cup \{\infty\}$. Quand un client arrive, il doit se faire servir par

- un vendeur (temps de service de loi exponentielle de paramètre $\mu_1 > 0$), puis
- par un caissier (temps de service de loi exponentielle de paramètre $\mu_2 > 0$), puis
- par un manutentionnaire (temps de service de loi exponentielle de paramètre $\mu_3 > 0$).

Les objets aléatoires ci-dessus sont bien sûr mutuellement indépendants. On a donc en fait trois files d'attentes qui interagissent les unes avec les autres.

Soit $X_t = (X_t^1, X_t^2, X_t^3)$ où, à l'instant $t \geq 0$,

- X_t^1 est le nombre de clients en train d'attendre ou de se faire servir par un vendeur,
- X_t^2 est le nombre de clients en train d'attendre ou de se faire servir par un caissier,
- X_t^3 est le nombre de clients en train d'attendre ou de se faire servir par un manutentionnaire.

Théorème 6.4. Le processus $(X_t)_{t \geq 0}$ est un PMS à valeurs dans $E = \mathbb{N}^3$. De plus, on a, pour tout $i = (i_1, i_2, i_3)$,

$$\lambda(i) = \alpha + \min\{i_1, s_1\}\mu_1 + \min\{i_2, s_2\}\mu_2 + \min\{i_3, s_3\}\mu_3,$$

et, pour $i = (i_1, i_2, i_3) \in \mathbb{N}^3$ et $j = (j_1, j_2, j_3) \in \mathbb{N}^3$,

$$Q(i, j) = \begin{cases} \frac{\alpha}{\lambda(i)} & \text{si } j = (i_1 + 1, i_2, i_3), \\ \frac{\min\{i_1, s_1\}\mu_1}{\lambda(i)} & \text{si } j = (i_1 - 1, i_2 + 1, i_3), \\ \frac{\min\{i_2, s_2\}\mu_2}{\lambda(i)} & \text{si } j = (i_1, i_2 - 1, i_3 + 1), \\ \frac{\min\{i_3, s_3\}\mu_3}{\lambda(i)} & \text{si } j = (i_1, i_2, i_3 - 1). \end{cases}$$

Dans tous les autres cas, on a $Q(i, j) = 0$.

Preuve. Etape 1. C'est un PMS car les lois en jeu sont toutes exponentielles (sans mémoire). Si on connaît $(X_s)_{s \in [0, t]}$, alors ce qui se passe après t ne dépend que de X_t . En effet, supposons que $X_t = i = (i_1, i_2, i_3)$. Alors :

- on sait qu'il y a un client en train d'arriver, et on sait depuis combien de temps on l'attend, mais, la loi des interarrivées $\text{Exp}(\alpha)$ étant sans mémoire, notre temps résiduel d'attente sera encore de loi $\text{Exp}(\alpha)$.
- D'autre part, il y a $\min\{i_1, s_1\}$ clients en train de se faire servir par des vendeurs, on sait depuis combien de temps, etc. Mais, les temps de service étant $\text{Exp}(\mu_1)$, les temps résiduels de service, à l'instant t , sont encore de loi $\text{Exp}(\mu_1)$.
- Il y a aussi $\min\{i_2, s_2\}$ clients en train de se faire servir par un caissier, on sait depuis combien de temps, etc. Mais, les temps de service étant $\text{Exp}(\mu_2)$, les temps résiduels de service, à l'instant t , sont encore de loi $\text{Exp}(\mu_2)$.
- Enfin, il y a $\min\{i_3, s_3\}$ clients en train de se faire servir par un manutentionnaire, on sait depuis combien de temps, etc. Mais, les temps de service étant $\text{Exp}(\mu_3)$, les temps résiduels de service, à l'instant t , sont encore de loi $\text{Exp}(\mu_3)$.

Ainsi, la connaissance de $(X_s)_{s \in [0, t]}$ n'apporte rien de plus que celle de X_t pour le comportement à venir.

Etape 2. On détermine le générateur. On pose donc $T_1 = \inf\{t \geq 0 : X_t \neq X_0\}$, et on cherche $Q(i, j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i[T_1]$. On suppose donc que $X_0 = i = (i_1, i_2, i_3)$. On pose $k_1 = \min\{i_1, s_1\}$, $k_2 = \min\{i_2, s_2\}$, $k_3 = \min\{i_3, s_3\}$.

- Soit $T \sim \text{Exp}(\alpha)$ le temps d'arrivée du prochain client.
- Soient $S_1^1, \dots, S_{k_1}^1 \sim \text{Exp}(\mu_1)$ les temps de service des clients en train de se faire servir par un vendeur. On note $U_1 = \min\{S_1^1, \dots, S_{k_1}^1\}$, qui suit la loi $\text{Exp}(k_1\mu_1)$ par le lemme des réveils (si $k_1 = 0$, ça donne $U_1 \sim \text{Exp}(0)$, qui vaut l'infini p.s. par convention).
- Soient $S_1^2, \dots, S_{k_2}^2 \sim \text{Exp}(\mu_2)$ les temps de service des clients en train de se faire servir par un caissier. On note $U_2 = \min\{S_1^2, \dots, S_{k_2}^2\}$, qui suit la loi $\text{Exp}(k_2\mu_2)$ par le lemme des réveils.
- Soient $S_1^3, \dots, S_{k_3}^3 \sim \text{Exp}(\mu_3)$ les temps de service des clients en train de se faire servir par un manutentionnaire. On note $U_3 = \min\{S_1^3, \dots, S_{k_3}^3\}$, qui suit la loi $\text{Exp}(k_3\mu_3)$ par le lemme des réveils.

On a donc $T_1 = \min\{T, U_1, U_2, U_3\}$ et $X_{T_1} = (i_1 + 1, i_2, i_3)$ si $T_1 = T$, $X_{T_1} = (i_1 - 1, i_2 + 1, i_3)$ si $T_1 = U_1$, $X_{T_1} = (i_1, i_2 - 1, i_3 + 1)$ si $T_1 = U_2$, $X_{T_1} = (i_1, i_2, i_3 - 1)$ si $T_1 = U_3$. Par le lemme des réveils, on conclut qu'en effet,

on a bien $T_1 \sim \mathcal{E}(\alpha + k_1\mu_1 + k_2\mu_2 + k_3\mu_3)$, donc

$$\lambda(i) = \alpha + k_1\mu_1 + k_2\mu_2 + k_3\mu_3,$$

et que

$$Q(i, (i_1 + 1, i_2, i_3)) = \mathbf{P}_i(X_{T_1} = (i_1 + 1, i_2, i_3)) = \mathbf{P}(T_1 = T) = \frac{\alpha}{\lambda(i)},$$

et

$$Q(i, (i_1 - 1, i_2 + 1, i_3)) = \mathbf{P}_i(X_{T_1} = (i_1 - 1, i_2 + 1, i_3)) = \mathbf{P}(T_1 = U_1) = \frac{k_1\mu_1}{\lambda(i)}.$$

Les autres valeurs s'obtiennent de manière similaire. $\square$

En utilisant la formule (6.1) avec la fonction $\phi(i_1, i_2, i_3) = i_1$ (puis $\phi(i_1, i_2, i_3) = i_2$ puis $\phi(i_1, i_2, i_3) = i_3$), on trouve

$$\begin{aligned} \frac{d}{dt} \mathbf{E}_i[X_t^1] &= \alpha - \mu_1 \mathbf{E}_i[\min\{X_t^1, s_1\}], \\ \frac{d}{dt} \mathbf{E}_i[X_t^2] &= \mu_1 \mathbf{E}_i[\min\{X_t^1, s_1\}] - \mu_2 \mathbf{E}_i[\min\{X_t^2, s_2\}], \\ \frac{d}{dt} \mathbf{E}_i[X_t^3] &= \mu_2 \mathbf{E}_i[\min\{X_t^2, s_2\}] - \mu_3 \mathbf{E}_i[\min\{X_t^3, s_3\}]. \end{aligned}$$

On peut certainement résoudre ce système lorsque $s_1 = s_2 = s_3 = \infty$.

6.6 Un modèle simple de populations

Considérons une population d'individus assexués. La durée de vie de chaque individu suit une loi exponentielle de paramètre μ . D'autre part, chaque individu (vivant) produit des portées suivant un processus de Poisson de paramètre α . La loi du nombre d'individus issus d'une portée est donnée par une probabilité ν sur $\mathbb{N}^*$. Enfin, des individus immigreront suivant un processus de Poisson de paramètre κ .

Soit X_t le nombre d'individus vivant à l'instant $t \geq 0$.

Théorème 6.5. Le processus $(X_t)_{t \geq 0}$ est un PMS à valeurs dans $E = \mathbb{N}$. De plus, on a, pour tout $i \in \mathbb{N}$,

$$\lambda(i) = \kappa + i\mu + i\alpha$$

et, pour $i = \mathbb{N}$ et $j \in \mathbb{N}$,

$$Q(i, j) = \begin{cases} \frac{i\mu}{\lambda(i)} & \text{si } j = i - 1, \\ \frac{\kappa + i\alpha\nu(1)}{\lambda(i)} & \text{si } j = i + 1, \\ \frac{i\alpha\nu(k)}{\lambda(i)} & \text{si } j = i + k \text{ (avec } k \geq 2), \end{cases}$$

et $Q(i, j) = 0$ dans tous les autres cas.

Preuve. Etape 1. C'est un PMS car les lois en jeu sont toutes exponentielles (sans mémoire). Si on connaît $(X_s)_{s \in [0, t]}$, alors ce qui se passe après t ne dépend que de X_t . En effet, supposons que $X_t = i$.

- Alors on sait qu'un immigrant arrive, et on sait depuis combien de temps on l'attend, mais, la loi des interarrivées $\text{Exp}(\kappa)$ étant sans mémoire, notre temps résiduel d'attente sera encore $\text{Exp}(\kappa)$.
- D'autre part, il y a i individus en train de mourir, et on sait quand ils sont nés, mais, comme leur durée de vie suit une loi exponentielle, leur durée de vie résiduelle à l'instant t est encore $\text{Exp}(\mu)$.
- Enfin, i individus vont produire (potentiellement) une portée. Pour chacun, on sait son âge et la date de sa dernière portée éventuelle, mais, les portées arrivant suivant des processus de Poisson de paramètre α , sa prochaine portée se produira dans un temps de loi $\text{Exp}(\alpha)$. Enfin, la taille de cette portée ne dépend de rien.

Ainsi, la connaissance de $(X_s)_{s \in [0, t]}$ n'apporte rien de plus que celle de X_t pour le comportement à venir.

Etape 2. On détermine le générateur. On pose donc $T_1 = \inf\{t \geq 0 : X_t \neq X_0\}$, et on cherche $Q(i, j) = \mathbf{P}_i(X_{T_1} = j)$ et $\lambda(i) = 1/\mathbf{E}_i[T_1]$. On suppose donc que $X_0 = i$.

- Soit $T \sim \text{Exp}(\kappa)$ le temps d'arrivée du prochain immigrant.
- Soient $M_1, \dots, M_i$ les dates de mort des i individus, on sait qu'elles sont i.i.d. de loi $\mathcal{E}(\mu)$.
- Soient $P_1, \dots, P_i$ les dates des premières portées des i individus, on sait qu'elles sont i.i.d. de loi $\text{Exp}(\alpha)$. Enfin, soient $Z_1, \dots, Z_i$ les tailles de ces portées, qui sont i.i.d. de loi ν .

On a donc $T_1 = \min\{T, M_1, \dots, M_i, P_1, \dots, P_i\}$ et par le lemme des $2i + 1$ réveils, on a donc $T_1 \sim \text{Exp}(\kappa + i\mu + i\alpha)$, soit $\lambda(i) = \kappa + i\mu + i\alpha$.

De plus, $X_{T_1} = i + 1$ si $T_1 = T$, $X_{T_1} = i - 1$ si $T_1 \in \{M_1, \dots, M_i\}$, et enfin, pour tout $j = 1, \dots, i$, $X_{T_1} = i + Z_j$ si $T_1 = P_j$.

Donc $X_{T_1} = i - 1$ si $T_1 \in \{M_1, \dots, M_i\}$, et par le lemme des réveils,

$$Q(i, i - 1) = \sum_{j=1}^i \mathbf{P}(T_1 = M_j) = \sum_{j=1}^i \frac{\mu}{\lambda(i)} = \frac{i\mu}{\lambda(i)}.$$

Aussi, $X_{T_1} = i + 1$ si $T_1 = T$ ou si, pour un $j \in \{1, \dots, i\}$, $T_1 = P_j$ et $Z_j = 1$, donc

$$Q(i, i + 1) = \mathbf{P}(T_1 = T) + \sum_{j=1}^i \mathbf{P}(T_1 = P_j, Z_j = 1) = \frac{\kappa}{\lambda(i)} + \sum_{j=1}^i \frac{\alpha}{\lambda(i)} \nu(1) = \frac{\kappa + i\alpha\nu(1)}{\lambda(i)}.$$

Enfin, pour $k \geq 2$, $X_{T_1} = i + k$ si, pour un $j \in \{1, \dots, i\}$, $T_1 = P_j$ et $Z_j = k$, donc

$$Q(i, i + k) = \sum_{j=1}^i \mathbf{P}(T_1 = P_j, Z_j = k) = \sum_{j=1}^i \frac{\alpha}{\lambda(i)} \nu(k) = \frac{i\alpha\nu(k)}{\lambda(i)}.$$

□

Kolmogorov forward. Vérifier que KF s'écrit, pour ce processus, pour tout $i \geq 0$, tout $j \geq 1$ (inutile d'écrire $P'_t(i, 0)$)

$$\begin{aligned} P'_t(i, j) = & -(\kappa + j\mu + j\alpha)P_t(i, j) + \mu(j + 1)P_t(i, j + 1) + \kappa P_t(i, j - 1) \\ & + \sum_{k=1}^j \alpha(j - k)\nu(k)P_t(i, j - k). \end{aligned}$$

Dans ce cadre, la formule (6.1) s'écrit

$$\begin{aligned}
\frac{d}{dt} \mathbf{E}_i[\phi(X_t)] &= \mathbf{E}_i \left[\lambda(X_t) \sum_{j \in E} [\phi(j) - \phi(X_t)] Q(X_t, j) \right] \\
&= \mathbf{E}_i \left[\lambda(X_t) [\phi(X_t - 1) - \phi(X_t)] \frac{X_t \mu}{\lambda(X_t)} \right] \\
&\quad + \mathbf{E}_i \left[\lambda(X_t) [\phi(X_t + 1) - \phi(X_t)] \frac{\kappa + X_t \alpha \nu(1)}{\lambda(X_t)} \right] \\
&\quad + \sum_{k \geq 2} \mathbf{E}_i \left[\lambda(X_t) [\phi(X_t + k) - \phi(X_t)] \frac{X_t \alpha \nu(k)}{\lambda(X_t)} \right] \\
&= \mu \mathbf{E}_i [\phi(X_t - 1) - \phi(X_t)] X_t \\
&\quad + \kappa \mathbf{E}_i [\phi(X_t + 1) - \phi(X_t)] \\
&\quad + \alpha \sum_{k \geq 1} \mathbf{E}_i [\phi(X_t + k) - \phi(X_t)] X_t \nu(k).
\end{aligned}$$

Avec $\phi(j) = j$, on trouve, en posant $\rho = \sum_{k \geq 1} k \nu(k)$,

$$\frac{d}{dt} \mathbf{E}_i[X_t] = -\mu \mathbf{E}_i[X_t] + \kappa + \alpha \rho \mathbf{E}_i[X_t] = \kappa + (\alpha \rho - \mu) \mathbf{E}_i[X_t].$$

Comme $\mathbf{E}_i[X_0] = i$, on trouve, si $\alpha \rho \neq \mu$ (sinon, c'est une autre formule explicite)

$$\mathbf{E}_i[X_t] = -\frac{\kappa}{\alpha \rho - \mu} + \left(i + \frac{\kappa}{\alpha \rho - \mu} \right) e^{(\alpha \rho - \mu)t}.$$

On trouve donc deux comportements bien différents :

- si $\alpha \rho < \mu$, alors $\lim_{t \rightarrow \infty} \mathbf{E}_i[X_t] = \kappa / (\mu - \alpha \rho)$, la population reste contrôlée pas,
- si $\alpha \rho > \mu$, alors $\lim_{t \rightarrow \infty} \mathbf{E}_i[X_t] = \infty$.

C'est logique, car $\alpha \rho$ est le taux de naissance des portées multiplié par le nombre moyen de bébés par portée, qu'on compare au taux μ de décès.

6.7 L'importance des lois exponentielles

Considérons une population de bactéries, composée initialement (à $t = 0$) d'un unique individu. Chaque bactérie se coupe en deux bactéries filles, au bout d'un temps de loi uniforme sur $[0, 1]$, tout étant indépendant de tout au maximum. On appelle X_t le nombre de bactéries à l'instant t . Bien sûr, si on remplaçait la loi uniforme par une loi exponentielle, ce processus serait un PMS. Montrons qu'avec la loi uniforme, ce n'est pas le cas.

On a

$$p := \mathbf{P}(X_{1.6} = 2 \mid X_{0.5} = 2, X_1 = 2) \neq \mathbf{P}(X_{1.6} = 2 \mid X_1 = 2) =: q.$$

En effet, soit U le temps de vie de l'ancêtre, et V, W les temps de vie de ses deux filles. Les v.a. U, V, W sont i.i.d. de loi uniforme sur $[0, 1]$.

$$p = \mathbf{P}(U + V > 1.6, U + W > 1.6 \mid U \leq 0.5, U + V > 0.5, U + W > 0.5) = 0,$$

car $U \leq 0.5$ implique $U + V \leq 1.5$ et $U + W \leq 1.5$.

$$\begin{aligned}
q &= \mathbf{P}(U + V > 1.6, U + W > 1.6 \mid U \leq 1, U + V > 1, U + W > 1) \\
&= \frac{\mathbf{P}(U + V > 1.6, U + W > 1.6)}{\mathbf{P}(U + V > 1, U + W > 1)} =: \frac{a}{b} > 0,
\end{aligned}$$

car

$$\begin{aligned} a &= \int_0^1 \int_0^1 \int_0^1 \mathbf{1}_{u+v>1.6, u+w>1.6} dw dv du = \int_{0.6}^1 du \int_{1.6-u}^1 dv \int_{1.6-u}^1 dw \\ &= \int_{0.6}^1 (u-0.6)^2 du = \frac{(0.4)^3}{3}, \end{aligned}$$

et

$$b = \int_0^1 \int_0^1 \int_0^1 \mathbf{1}_{u+v>1, u+w>1} dw dv du = \int_0^1 du \int_{1-u}^1 dv \int_{1-u}^1 dw = \int_0^1 u^2 du = \frac{1}{3}.$$

A

Compléments

A.1 π -systèmes et classes monotones

Ce paragraphe contient quelques rappels de théorie de la mesure particulièrement utiles lorsqu'il s'agit de caractériser les lois des variables aléatoires. Ils permettent notamment de montrer qu'il suffit souvent de ne considérer qu'une sous-classe d'événements ou de fonctions.

Définition (π -système). Une classe $\mathcal{P}$ de parties d'un ensemble S est un π -système si elle est stable par intersection: pour tout $A, B \in \mathcal{P}$ on a $A \cap B \in \mathcal{P}$.

Exemple A.1. Soit $\mathcal{P}$ la collection des intervalles ouverts (a, b) dans $\mathbb{R}$, potentiellement $(a, a) = \emptyset$. Alors, $\mathcal{P}$ est un π -système puisque l'intersection de deux intervalles ouverts est soit un intervalle ouvert non-trivial, soit l'ensemble vide.

Définition (Classe monotone / Dynkin-système). Une classe $\mathcal{M}$ de parties d'un ensemble S est une classe monotone si

- i) $S \in \mathcal{M}$;
- ii) si $A, B \in \mathcal{M}$ et $A \subseteq B$ alors $B \setminus A \in \mathcal{M}$;
- iii) si $A_n \in \mathcal{M}$ pour tout $n \geq 0$ et $A_n \uparrow A$ (donc $A_n \subseteq A_{n+1}$ pour tout n) alors $A \in \mathcal{M}$.

Lemme A.1 (Classes monotones – version ensembliste). Soit Π un π -système et $\mathcal{M}$ une classe monotone qui contient Π . Alors $\sigma(\Pi) \subseteq \mathcal{M}$.

Le lemme des classes monotones est typiquement utilisé pour montrer que deux mesures sont égales: on montre qu'elles coïncident sur un π -système.

Corollaire A.2. Soit $(\Omega, \mathcal{F})$ un espace mesurable. Soit $\mathcal{P}$ un π -système de parties de Ω qui engendre $\mathcal{F}$. Si deux mesures sur $(E, \mathcal{F})$ coïncident sur tout élément de $\mathcal{P}$, alors elles coïncident sur $\sigma(\mathcal{P}) = \mathcal{F}$.

Exemple A.2. On considère $\mathbb{R}$ muni de la tribu borélienne $\mathcal{B}$. Alors le π -système $\mathcal{P} = \{(a, b) : a \leq b\}$ engendre $\mathcal{B}$. Deux mesures de probabilité μ et ν qui coïncident sur $\mathcal{P}$ sont identiques. En effet, la collection $\mathcal{M}$ de parties de $\mathbb{R}$ définie par $\mathcal{M} := \{A \in \mathcal{B} : \mu(A) = \nu(A)\}$ est une classe monotone:

- $\mu(\mathbb{R}) = 1 = \nu(\mathbb{R})$ donc $\mathbb{R} \in \mathcal{M}$;
- si $A, B \in \mathcal{M}$ et $A \subseteq B$ alors $\mu(B \setminus A) = \mu(B) - \mu(A) = \nu(B) - \nu(A) = \nu(B \setminus A)$ donc $B \setminus A \in \mathcal{M}$;
- si $A_n \in \mathcal{M}$ pour tout $n \in \mathbb{N}$ et $A_n \uparrow A$ alors $\mu(A) = \lim_n \mu(A_n) = \lim_n \nu(A_n) = \nu(A)$.

Comme $\mathcal{P} \subseteq \mathcal{M}$ par hypothèse, on en déduit que $\mathcal{B} = \sigma(\mathcal{P}) \subseteq \mathcal{M}$, et donc que μ et ν sont égales.

Il est parfois utile d'avoir sous la main une version fonctionnelle: au lieu de considérer des classes de parties d'ensemble et d'essayer de montrer que cette classe contient une certaine tribu, on considère une classe de fonctions et l'objectif est de montrer qu'elle contient toutes les fonctions mesurables.

Lemme A.3 (Classes monotones – version fonctionnelle). Soit $\mathcal{H}$ une collection de fonctions bornées d'un ensemble S dans $\mathbb{R}$ qui satisfait les conditions suivantes:

- $\mathcal{H}$ est un espace vectoriel: pour tout $\lambda, \mu \in \mathbb{R}$ et $f, g \in \mathcal{H}$ on a $\lambda f + \mu g \in \mathcal{H}$;
- la fonction constante égale à un est un élément de $\mathcal{H}$;
- si $(f_n)_{n \geq 0}$ est une suite de fonctions positives dans $\mathcal{H}$ telle que $f_n \uparrow f$ où f est une fonction bornée sur S alors $f \in \mathcal{H}$.

On suppose additionnellement que $\mathcal{H}$ contient les fonctions indicatrices des ensembles dans un π -système $\mathcal{P}$. Alors, $\mathcal{H}$ contient toutes les fonctions bornées $\sigma(\mathcal{P})$ -mesurables sur S .

A.2 A propos des lois des processus stochastiques

On considère un espace mesurable $(\Omega, \mathcal{F}, \mathbf{P})$. Un processus stochastique est une famille de variables aléatoires $(X_t)_{t \in T}$ à valeurs dans un espace mesurable $(E, \mathcal{E})$, avec $T = \mathbb{N}$ ou $\mathbb{R}_+$ (parfois $\mathbb{Z}$ ou $\mathbb{R}$). On peut considérer $X = (X_t)_{t \in T}$ comme une variable aléatoire à valeur dans E^T , pour peu qu'on munisse cet espace d'une tribu adaptée. La tribu naturelle est la tribu produit $\mathcal{E}^{\otimes T}$, c'est-à-dire la tribu sur E^T engendrée par les cylindres $\{\mathbf{x} \in E^T : x_t \in B_t, t \in S\}$ pour $S \subseteq T$ un ensemble fini et $B_t \in \mathcal{E}$ pour tout $t \in S$.

Soit $C = \{\mathbf{x} \in E^T : x_t \in B_t, t \in S\}$ un cylindre alors

$$\{X \in C\} = \{(X_t)_{t \in T} \in C\} = \bigcap_{t \in S} \{X_t \in B_t\} \in \mathcal{F}$$

puisque toutes les X_t sont des variables aléatoires et $S \subseteq T$ est fini.

Par ailleurs, on voit aisément que la collection $\mathcal{C}$ des cylindres sur E^T est un π -système; ce π -système engendre $\mathcal{E}^T$ par définition. Soit $\mathcal{M}$ la collection de parties de $\mathcal{E}^T$ définie par

$$\mathcal{M} := \left\{ A \in \mathcal{E}^T : \{X \in A\} \in \mathcal{F} \right\}.$$

Nous venons de montrer que $\mathcal{M}$ contient l'ensemble des cylindres. De plus, $\mathcal{M}$ est une classe monotone puisque

- $E^T \in \mathcal{M}$ car $\{X \in E^T\} = \Omega \in \mathcal{F}$;
- si $A, B \in \mathcal{M}$ et $A \subseteq B$ alors $\{X \in B \setminus A\} = \{X \in B\} \setminus \{X \in A\} \in \mathcal{F}$;
- si $A_n, n \in \mathbb{N}$, et $A_n \uparrow A$ alors $\{X \in A\} = \bigcup_{n \geq 0} \{X \in A_n\} \in \mathcal{F}$.

On déduit du lemme des classes monotones (Lemma A.1) que $\mathcal{E}^T = \sigma(\mathcal{C}) \subseteq \mathcal{M}$, ce qui signifie que la fonction $X : (\Omega, \mathcal{F}) \rightarrow (E^T, \mathcal{E}^T)$ qui à $\omega \in \Omega$ associe $X(\omega) = (X_t(\omega))_{t \in T}$ est mesurable.

Ceci justifie la “définition” d’égalité en loi pour les processus stochastiques (Définition). En effet, si pour tout $n \in \mathbb{N}$ et pour tous $t_0 < t_1 < \dots < t_n$ les vecteurs

$$(X_{t_1}, X_{t_2}, \dots, X_{t_n}) \quad \text{et} \quad (Y_{t_1}, Y_{t_2}, \dots, Y_{t_n})$$

ont même loi, alors les mesures μ_X et μ_Y sur $(E^T, \mathcal{E}^T)$ coïncident sur les cylindres: pour un cylindre $C = \{\mathbf{x} \in E^T : x_{t_i} \in B_i, 1 \leq i \leq n\}$ avec $B_1, B_2, \dots, B_n \in \mathcal{E}$ on a

$$\mu_X(C) = \mathbf{P}((X_{t_1}, \dots, X_{t_n}) \in B_1 \times B_2 \times \dots \times B_n) = \mathbf{P}((Y_{t_1}, \dots, Y_{t_n}) \in B_1 \times B_2 \times \dots \times B_n) = \mu_Y(C).$$

Comme l’ensemble des cylindres est un π -système qui engendre $\mathcal{E}^T$, les deux mesures μ_X et μ_Y sont en fait égales, et les processus X et Y ont donc même loi.

Remarque (Tribu produit et temps discret). L’ensemble E^T et sa tribu produit $\mathcal{E}^T$ ne sont en fait adéquats qu’en temps discret, c’est-à-dire dans le cas où T est dénombrable. En effet, dans le cas des processus à temps discret, on peut contrôler toute la trajectoire. En effet, pour toute suite $(B_i)_{i \in \mathbb{N}}$ d’éléments de $\mathcal{E}$, on a

$$\bigtimes_{i \in \mathbb{N}} B_i = \{\mathbf{x} \in E^{\mathbb{N}} : x_i \in B_i, i \geq 0\} = \bigcap_{k \in \mathbb{N}} B_1 \times B_2 \times \dots \times B_k \times E \times \dots \in \mathcal{E}^{\otimes \mathbb{N}}.$$

Remarque (Tribu produit et temps continu). Pour les processus en temps continu, c’est-à-dire lorsque T n’est plus dénombrable (typiquement, $T = \mathbb{R}_+$), on ne peut plus contrôler complètement une trajectoire (sans hypothèse supplémentaire) : intuitivement il n’est possible de contrôler les trajectoires qu’en un nombre dénombrable de positions. Un ensemble B est un élément de $\mathcal{E}^{\mathbb{R}_+}$ si et seulement si c’est un σ -cylindre, c’est-à-dire s’il s’écrit comme $\{\mathbf{x} \in E^{\mathbb{R}_+} : (x_t : t \in S) \in A_S\}$ pour un ensemble dénombrable $S \subseteq \mathbb{R}_+$ et $A_S \in \mathcal{E}^{\otimes S}$. Pour vérifier ceci, il suffit de montrer que l’ensemble des σ -cylindres est une tribu (ça sera alors une tribu qui contient les cylindres, qui contiendra donc la tribu produit $\mathcal{E}^T$; réciproquement, tout σ -cylindre est dans $\mathcal{E}^T$ comme intersection dénombrable de cylindres).

C’est pourquoi on travaille avec des processus continus à droite : pour un processus continu à droite (ou continu), il suffit de contrôler le processus à des positions denses dans $\mathbb{R}$ pour contrôler toute la trajectoire. On remarquera quand même la difficulté suivante, qui impose de travailler avec des processus qui sont *a priori* continus ou continus à droite: soit X un processus défini sur $(E^{\mathbb{R}_+}, \mathcal{E}^{\otimes \mathbb{R}_+})$, alors $\{X \text{ est continu}\}$ et $\{X \text{ est continu à droite}\}$ ne sont pas des ensembles mesurables, ni même des choses aussi naturelles que $\sup\{X_s : s \in [0, 1]\}$ (voir la fin de la remarque précédente).

Remarque. Même en temps continu, l’ensemble des fonctions $f : E^{\mathbb{R}_+} \rightarrow \mathbb{R}$ mesurables est engendré par les indicatrices des cylindres (lemme des classes monotones, version fonctionnelle). C’est pourquoi, pour caractériser la loi d’un processus $X = (X_t)_{t \geq 0}$ il suffit de calculer $\mathbf{E}[f(X)]$ pour les fonctions $f : E^{\mathbb{R}_+} \rightarrow \mathbb{R}_+$ de la forme $\mathbf{x} = (x_t)_{t \geq 0} \in E^{\mathbb{R}_+} \mapsto f(\mathbf{x}) = \phi(\pi_{t_1}(\mathbf{x}), \dots, \pi_{t_n}(\mathbf{x}))$ avec $\phi : E^n \rightarrow \mathbb{R}_+$ et $\pi_t(\mathbf{x}) = x_t$. C’est par exemple utilisé dans la preuve du théorème 5.3.

A.3 Propriété de Markov pour les trajectoires des chaînes de Markov

On justifie ici proprement les calculs d'espérance conditionnelles qui impliquent la loi des trajectoires de chaînes de Markov (preuve du Théorème 4.5). Il s'agit de montrer que pour toute fonction $f : E^{\mathbb{N}} \rightarrow \mathbb{R}_+$ mesurable, on a

$$\mathbf{E}_v[f((X_{n+i})_{i \geq 0}) \mid \mathcal{F}_n] = \mathbf{E}_{X_n}[f((X_i)_{i \geq 0})], \quad (E)$$

Dans un premier temps, nous allons montrer qu'il suffit de montrer (E) pour une classe restreinte de fonctions f en utilisant les classes monotones.

Soit $\Pi = \{A = A_0 \times A_1 \times \cdots \times A_d, d \in \mathbb{N}, A_0, A_1, \dots, A_d \in \mathcal{E}\}$. Alors Π est un π -système de parties de $\mathcal{E}^{\mathbb{N}}$: pour $A, A' \in \Pi$ on a peut supposer que $d' = d$, quitte à prendre certains des ensembles A'_i ou A_i égaux à E , et alors

$$A \cap A' = (A_0 \cap A'_0) \times (A_1 \cap A'_1) \times \cdots \times (A_d \cap A'_d) \times E \times \cdots \in \Pi.$$

De plus, l'ensemble $\mathcal{M} \subseteq \mathcal{E}^{\otimes \mathbb{N}}$ défini par

$$\mathcal{M} = \{A \in \mathcal{E}^{\otimes \mathbb{N}} : \mathbf{E}_v[\mathbf{1}_A(X_n, X_{n+1}, \dots) \mid \mathcal{F}_n] = \mathbf{E}_{X_n}[\mathbf{1}_A(X_0, X_1, \dots)]\}$$

est une classe monotone. En effet:

- $E^{\mathbb{N}} \in \mathcal{M}$, puisqu'alors les deux membres de l'égalité dans la définition de $\mathcal{M}$ valent alors 1;
- si $A, B \in \mathcal{M}$ avec $A \subseteq B$ alors $B \setminus A \in \mathcal{M}$ puisque $\mathbf{1}_{B \setminus A} = \mathbf{1}_B - \mathbf{1}_A$;
- si $A_i \in \mathcal{M}$ pour tout $i \geq \mathbb{N}$ et $A_n \uparrow A$ alors $\mathbf{1}_{A_i} \uparrow \mathbf{1}_A$, et donc par convergence monotone, on voit aisément que $A \in \mathcal{M}$:

$$\begin{aligned} \mathbf{E}_v[\mathbf{1}_A(X_n, X_{n+1}, \dots) \mid \mathcal{F}_n] &= \lim_i \mathbf{E}_v[\mathbf{1}_{A_i}(X_n, X_{n+1}, \dots) \mid \mathcal{F}_n] \\ &= \lim_i \mathbf{E}_{X_n}[\mathbf{1}_{A_i}(X_0, X_1, \dots)] = \mathbf{E}_{X_n}[\mathbf{1}_A(X_0, X_1, \dots)]. \end{aligned}$$

Par conséquent, d'après le lemme des classes monotones (version fonctionnelle), pour montrer que (E) est satisfaite pour toute fonction $f : E^{\mathbb{N}} \rightarrow \mathbb{R}_+$ mesurable bornée, il suffit de le vérifier pour la collection de fonctions $\{\mathbf{1}_A, A \in \Pi\}$: en effet, si on note $\mathcal{H}$ l'ensemble des fonctions mesurables bornées telles que l'identité (E) est vérifiée alors

- $\mathcal{H}$ est clairement un espace vectoriel par linéarité de l'espérance;
- $\mathcal{H}$ contient la fonction constante égale à un;
- par convergence monotone si $f_i \in \mathcal{H}$ pour tout $i \in \mathbb{N}$ et $f_i \uparrow f$ avec f bornée, alors $f \in \mathcal{H}$.

Maintenant, pour montrer que, pour tout $A \in \Pi$, on a

$$\mathbf{E}_v[\mathbf{1}_A(X_n, X_{n+1}, \dots) \mid \mathcal{F}_n] = \mathbf{E}_{X_n}[\mathbf{1}_A(X_0, X_1, \dots)],$$

il faut vérifier (1) que $\mathbf{E}_{X_n}[f((X_i)_{i \geq 0})]$ est $\mathcal{F}_n$ -mesurable, ce qui est clair car cette fonction est $\sigma(X_n)$ -mesurable, et (2) que pour tout $B \in \mathcal{F}_n$ on a

$$\mathbf{E}_v[\mathbf{E}_{X_n}[\mathbf{1}_A((X_i)_{i \geq 0})] \mathbf{1}_B] = \mathbf{E}_v[\mathbf{1}_A((X_{n+i})_{i \geq 0}) \mathbf{1}_B]. \quad (\star)$$

Hors, la tribu $\mathcal{F}_n$ est engendrée par les événements

$$\{X_0 \in C_0, X_1 \in C_1, \dots, X_n \in C_n\} = \bigcup_{x_0 \in C_0} \cdots \bigcup_{x_n \in C_n} \{X_0 = x_0, X_1 = x_1, \dots, X_n = x_n\}$$

avec $C_i \in \mathcal{E}$ pour tout $i = 0, 1, \dots, n$ qui forment un π -système. Pour chacun des ensembles $B = C_0 \times C_1 \times \dots \times C_n$, on a donc par linéarité, quelque soit la variable aléatoire Z ,

$$\mathbf{E}[Z \mathbf{1}_B] = \sum_{x_0 \in C_0} \dots \sum_{x_n \in C_n} \mathbf{E}[Z \mathbf{1}_{X_0=x_0, \dots, X_n=x_n}].$$

Pour montrer que $(\star)$ est valide pour tout $B \in \mathcal{F}_n$, il suffit donc

- de le montrer pour les événements $\{X_0 = x_0, X_1 = x_1, \dots, X_n = x_n\}$,
- d'étendre à tous les $\{X_0 \in C_0, X_1 \in C_1, \dots, X_n \in C_n\}$ par linéarité,
- d'étendre à tous les $B \in \mathcal{F}_n$ par classe monotone.

On est donc conduit à vérifier que pour tout $A \in \Pi$ et tous éléments $x_0, x_1, \dots, x_n \in E$ on a

$$\mathbf{E}_v[\mathbf{E}_{X_n}[\mathbf{1}_A((X_i)_{i \geq 0})] \mathbf{1}_{X_0=x_0, X_1=x_1, \dots, X_n=x_n}] = \mathbf{E}_v[\mathbf{1}_A((X_{n+i})_{i \geq 0}) \mathbf{1}_{X_0=x_0, X_1=x_1, \dots, X_n=x_n}]. \quad (\star\star)$$

Encore une fois, comme E est dénombrable, pour chaque élément $A = A_0 \times A_1 \times \dots \times A_d \times E^{\mathbb{N}} \in \mathcal{P}$, on a

$$\mathbf{1}_A = \sum_{a_0 \in A_0} \dots \sum_{a_d \in A_d} \mathbf{1}_{\{a_0\} \times \{a_1\} \times \dots \times \{a_d\} \times E \times \dots}.$$

Par conséquent, si on montre $(\star\star)$ pour les ensembles $A = \{a_0\} \times \{a_1\} \times \dots \times \{a_d\} \times E \times \dots$ alors, on l'aura aussi pour les ensembles $A_0 \times A_1 \times \dots \times A_d \times E \times \dots \subseteq E^{\mathbb{N}}$ par linéarité. C'est précisément ce calcul qui est fait dans la preuve du théorème puisqu'on y vérifie que les deux probabilités

$$\mathbf{P}_v(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, X_n = a_0, X_{n+1} = a_1, \dots, X_{n+d} = a_d)$$

et

$$\mathbf{P}_v(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) \cdot \mathbf{P}_{x_n}(X_0 = a_0, X_1 = a_1, \dots, X_d = a_d)$$

sont égales.

B

Appendice

Théorème B.1. Soit $(S_n)_{n \geq 0}$ une marche aléatoire à valeurs dans $\mathbb{Z}$, de pas de loi μ (une probabilité sur $\mathbb{Z}$). C'est bien sûr une chaîne de transition $P(i, j) = \mu(j - i)$. Si μ est centrée, i.e. si $\sum_{k \in \mathbb{Z}} k\mu(k) = 0$, alors tout état est récurrent.

Preuve. (a) On considère $(X_k)_{k \geq 1}$ i.i.d. de loi μ . Pour tout $i \in \mathbb{Z}$, $S_n = i + X_1 + \dots + X_n$ est une chaîne de transition P issue de i .

(b) Pour tous $i, j \in \mathbb{Z}$, on a $G(i, j) = G(0, j - i)$, car

$$G(i, j) = \sum_{n \geq 0} \mathbf{P}(i + X_1 + \dots + X_n = j) = \sum_{n \geq 0} \mathbf{P}(X_1 + \dots + X_n = j - i) = G(0, j - i).$$

(c) Donc pour tout $i \in \mathbb{Z}$, $G(i, i) = G(0, 0)$: il suffit de montrer que $G(0, 0) = \infty$.

(d) Par le principe du maximum, on a $G(0, 0) \geq G(-i, 0) = G(0, i)$ pour tout $i \in \mathbb{Z}$.

(e) Soient $L \in \mathbb{N}$ et $A \in \mathbb{N}$. Par (d), on a

$$\begin{aligned} G(0, 0) &\geq \frac{1}{2L+1} \sum_{i=-L}^L G(0, i) \\ &= \frac{1}{2L+1} \sum_{n \geq 0} \mathbf{P}(S_n \in \{-L, \dots, L\}) \\ &\geq \frac{1}{2L+1} \sum_{n=0}^{AL} \mathbf{P}\left(\frac{S_n}{n} \in \left\{-\frac{1}{A}, \dots, \frac{1}{A}\right\}\right) \end{aligned}$$

car $n \leq AL$ implique que $L/n \geq 1/A$.

(f) Soit $A \in \mathbb{N}$ grand. Par la LGN (faible) et comme $\mathbf{E}[X_1] = 0$, on sait que

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\frac{S_n}{n} \in \left\{-\frac{1}{A}, \dots, \frac{1}{A}\right\}\right) = 1.$$

Soit $n_A \in \mathbb{N}^*$ tel que pour tout $n \geq n_A$,

$$\mathbf{P}\left(\frac{S_n}{n} \in \left\{-\frac{1}{A}, \dots, \frac{1}{A}\right\}\right) \geq \frac{1}{2}.$$

Soit ensuite $L_A \in \mathbb{N}_*$ tel que $AL_A \geq 2n_A$ et AL_A soit pair (il suffit de poser $L_A = 2\lfloor n_A/A \rfloor + 2$). Alors

$$\begin{aligned}
 G(0,0) &\geq \frac{1}{2L_A+1} \sum_{n=0}^{AL_A} \mathbf{P}\left(\frac{S_n}{n} \in \left\{-\frac{1}{A}, \dots, \frac{1}{A}\right\}\right) \\
 &\geq \frac{1}{2L_A+1} \sum_{n=AL_A/2}^{AL_A} \mathbf{P}\left(\frac{S_n}{n} \in \left\{-\frac{1}{A}, \dots, \frac{1}{A}\right\}\right) \\
 &\geq \frac{1}{2L_A+1} \sum_{n=AL_A/2}^{AL_A} \frac{1}{2} \\
 &= \frac{1}{2L_A+1} \frac{AL_A}{4} \\
 &\geq \frac{1}{3L_A} \frac{AL_A}{4} = \frac{A}{12}.
 \end{aligned}$$

(On a bien $2L+1 \leq 3L$ pour tout entier L non nul). Ceci étant vrai pour tout $A \in \mathbb{N}$, $G(0,0) = \infty$. $\square$

Proposition B.2. Soit $s \in \mathbb{N} \cup \{\infty\}$. Considérons la file $M/M/s$, i.e. le PMS $(X_t)_{t \geq 0}$ à valeurs dans $\mathbb{N}$, avec $\lambda(i) = \lambda + (i \wedge s)\mu$, $Q(0,1) = 1$, et $Q(i, i-1) = (i \wedge s)\mu/\lambda(i)$, $Q(i, i+1) = \alpha/\lambda(i)$, quand $i \geq 1$. Alors $t \mapsto \mathbf{E}_i[X_t]$ est C^1 sur $\mathbb{R}_+$, et on a

$$\frac{d}{dt} \mathbf{E}_i[X_t] = \alpha - \mu \mathbf{E}_i[X_t \wedge s].$$

Preuve. (a) Soit $\phi : \mathbb{N} \rightarrow \mathbb{R}$ à support compact (il existe $A > 0$ tel que $\phi(i) = 0$ pour tout $i \geq A$). Alors la formule (6.1) est parfaitement justifiée (car on échange une somme finie et une dérivation), et on trouve

$$\frac{d}{dt} \mathbf{E}_i[\phi(X_t)] = \alpha \mathbf{E}_i[\phi(X_t+1) - \phi(X_t)] + \mu \mathbf{E}_i[(X_t \wedge s)(\phi(X_t-1) - \phi(X_t))]. \quad (\text{B.1})$$

(b) Soit $\phi : \mathbb{N} \rightarrow \mathbb{R}$ telle qu'il existe A, B tels que $\phi(i) = B$ pour tout $i \geq A$. Alors la formule (B.1) est encore valable : on introduit $\psi(i) = \phi(i) - B$, à laquelle on a le droit d'appliquer (B.1), et on trouve

$$\frac{d}{dt} \mathbf{E}_i[\psi(X_t)] = \alpha \mathbf{E}_i[\psi(X_t+1) - \psi(X_t)] + \mu \mathbf{E}_i[(X_t \wedge s)(\psi(X_t-1) - \psi(X_t))].$$

Mais comme $\mathbf{E}_i[\phi(X_t)] = \mathbf{E}_i[\psi(X_t)] + B$, on a $\frac{d}{dt} \mathbf{E}_i[\phi(X_t)] = \frac{d}{dt} \mathbf{E}_i[\psi(X_t)]$. Comme de plus $\phi(a) - \phi(b) = \psi(a) - \psi(b)$ pour tous a, b , on conclut aisément.

(c) On a donc le droit d'appliquer (B.1) avec $\phi(x) = x \wedge A$, pour $A > 0$ fixé :

$$\frac{d}{dt} \mathbf{E}_i[X_t \wedge A] = \alpha \mathbf{E}_i[(X_t+1) \wedge A - X_t \wedge A] + \mu \mathbf{E}_i[(X_t \wedge s)((X_t-1) \wedge A - X_t \wedge A)]. \quad (\text{B.2})$$

Le dernier terme est négatif, donc

$$\frac{d}{dt} \mathbf{E}_i[X_t \wedge A] \leq \alpha \mathbf{E}_i[(X_t+1) \wedge A - X_t \wedge A] \leq \alpha.$$

On conclut que $\mathbf{E}_i[X_t \wedge A] \leq i + \alpha t$ pour tout $i \in \mathbb{N}$, tout $t \geq 0$, tout $A > 0$. Par Fatou, on conclut que $\mathbf{E}_i[X_t] \leq i + \alpha t$ pour tout $i \in \mathbb{N}$, tout $t \geq 0$.

(d) On revient à (B.2), qu'on intègre en temps :

$$\begin{aligned} \mathbf{E}_i[X_t \wedge A] &= i \wedge A + \alpha \int_0^t \mathbf{E}_i[(X_u + 1) \wedge A - X_u \wedge A] du \\ &\quad + \mu \int_0^t \mathbf{E}_i[(X_u \wedge s)((X_u - 1) \wedge A - X_u \wedge A)] du. \end{aligned} \quad (\text{B.3})$$

On fait ensuite tendre $A \rightarrow \infty$ dans cette formule, par convergence dominée (pour $\mathbf{E}_i$ puis pour $\int_0^t$), en utilisant que pour tout $x \in \mathbb{N}$, tout $A > 0$,

- pour tout $x \in \mathbb{N}$, tout $A > 0$, $|(x + 1) \wedge A - x \wedge A| \leq 1$,
- pour tout $x \in \mathbb{N}$, tout $A > 0$, $(x \wedge s)((x - 1) \wedge A - x \wedge A) \leq x$,
- $\int_0^t \mathbf{E}_i[X_u] du < \infty$ puisque $0 \leq \mathbf{E}_i[X_t] \leq i + \alpha t$ par (c),
- pour tout $x \in \mathbb{N}$, $\lim_{A \rightarrow \infty} [(x + 1) \wedge A - x \wedge A] = 1$,
- pour tout $x \in \mathbb{N}$, $\lim_{A \rightarrow \infty} (x \wedge s)((x - 1) \wedge A - x \wedge A) = -(x \wedge s)$.

Ainsi, on a

$$\mathbf{E}_i[X_t] = i + \alpha t - \mu \int_0^t \mathbf{E}_i[X_u \wedge s] du.$$

(e) Si $s = \infty$, on a donc $\mathbf{E}_i[X_t] = i + \alpha t - \mu \int_0^t \mathbf{E}_i[X_u] du$. On déduit aisément de cette formule et du fait que $0 \leq \mathbf{E}_i[X_t] \leq i + \alpha t$ par (c), que $t \mapsto \mathbf{E}_i[X_t]$ est continue, puis C^1 . On dérive la formule, et on trouve que $\frac{d}{dt} \mathbf{E}_i[X_t] = \alpha - \mu \mathbf{E}_i[X_t]$.

Si $s \in \mathbb{N}$, on déduit de (B.3) avec $A = s$ que $t \mapsto \mathbf{E}_i[X_t \wedge s]$ est continue, ce qui implique que $t \mapsto \mathbf{E}_i[X_t] = i + \alpha t - \mu \int_0^t \mathbf{E}_i[X_u \wedge s] du$ est C^1 . On dérive la formule, et on trouve que $\frac{d}{dt} \mathbf{E}_i[X_t] = \alpha - \mu \mathbf{E}_i[X_t \wedge s]$. $\square$

Proposition B.3. On considère un PNM $(X_t)_{t \geq 0}$ irréductible. Si

$$R = \sum_{n=0}^{\infty} \left(\frac{1}{\beta_n} + \frac{\alpha_n}{\beta_n \beta_{n-1}} + \frac{\alpha_n \alpha_{n-1}}{\beta_n \beta_{n-1} \beta_{n-2}} + \cdots + \frac{\alpha_n \cdots \alpha_1}{\beta_n \cdots \beta_0} \right) = \infty,$$

alors il n'y a p.s. pas explosion : pour tout $i \in \mathbb{N}$, $P_i(\lim_n T_n < \infty) = 0$.

Preuve. On pose $u_i = \mathbf{E}_i[\exp(-T_\infty)]$, où $T_\infty = \lim_n T_n$. L'objectif est de montrer que $u_i = 0$ pour tout $i \in \mathbb{N}$.

Par Markov forte, en posant $\tilde{X}_t = X_{T_1+t}$, on a $T_\infty = T_1 + \tilde{T}_\infty$, donc

$$u_i = \mathbf{E}_i[\mathbf{E}_i[\exp(-T_1 - \tilde{T}_\infty) | \mathcal{F}_{T_1}]] = \mathbf{E}_i[\exp(-T_1) u_{X_1}] = \frac{\lambda(i)}{1 + \lambda(i)} \sum_{j \in \mathbb{N}} Q(i, j) u_j.$$

On a utilisé que si $T \sim \text{Exp}(\lambda)$, alors $\mathbf{E}[\exp(-T)] = \lambda / (1 + \lambda)$. On trouve, en remplaçant $\lambda(i)$ et $Q(i, j)$ par leurs valeurs,

$$u_0 = \frac{\beta_0}{1 + \beta_0} u_1, \quad u_i = \frac{\alpha_i u_{i-1} + \beta_i u_{i+1}}{1 + \alpha_i + \beta_i}, \quad i \geq 1.$$

Puis, en réorganisant,

$$u_1 - u_0 = \frac{u_0}{\beta_0}, \quad u_{i+1} - u_i = \frac{u_i}{\beta_i} + \frac{\alpha_i}{\beta_i} (u_i - u_{i-1}), \quad i \geq 1. \quad (\text{B.4})$$

Si $u_0 = 0$, on conclut facilement, par récurrence, que $u_i = 0$ pour tout $i \in \mathbb{N}$.

Reste à vérifier que $u_0 = 0$. On montre par récurrence, avec (B.4), que pour tout $i \geq 0$,

$$u_{i+1} - u_i = \frac{u_i}{\beta_i} + \sum_{k=0}^{i-1} \frac{\alpha_{k+1} \dots \alpha_i u_k}{\beta_k \dots \beta_i}.$$

Mais, toujours par (B.4) et par récurrence, on voit que pour tout $i \geq 0$, $u_{i+1} - u_i \geq 0$. Donc $u_k \geq u_0$ pour tout $k \geq 0$, et on conclut que pour tout $i \geq 0$,

$$u_{i+1} - u_i \geq u_0 \left(\frac{1}{\beta_i} + \sum_{k=0}^{i-1} \frac{\alpha_{k+1} \dots \alpha_i}{\beta_k \dots \beta_i} \right) = u_0 \left(\frac{1}{\beta_i} + \frac{\alpha_i}{\beta_i \beta_{i-1}} + \dots + \frac{\alpha_i \dots \alpha_1}{\beta_i \dots \beta_0} \right)$$

puis que

$$u_n = u_0 + \sum_{i=0}^{n-1} (u_{i+1} - u_i) \geq u_0 + u_0 \sum_{i=0}^{n-1} \left(\frac{1}{\beta_i} + \frac{\alpha_i}{\beta_i \beta_{i-1}} + \dots + \frac{\alpha_i \dots \alpha_1}{\beta_i \dots \beta_0} \right).$$

Comme $R = \infty$, on conclut que si $u_0 > 0$, alors $\lim_n u_n = \infty$, ce qui est absurde puisque $u_n \in [0, 1]$. Ainsi $u_0 = 0$. □